Copyright
by
Burcu Tan
2010

The Dissertation Committee for Burcu Tan
certifies that this is the approved version of the following dissertation:

Risk Mitigation Strategies for Project Management, Platform Development and Supply Chain Design

Committee:

Edward G. Anderson Jr., Supervisor

Qi (Annabelle) Feng

Geoffrey Parker

Jim S. Dyer

Sridhar Seshadri

Risk Mitigation Strategies for Project Management, Platform Development and Supply Chain Design

by

Burcu Tan, B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2010

Dedicated to my beloved husband Ozgur.

Acknowledgments

Many people have helped me in the process of writing this dissertation and I would like to express my gratitude to these individuals for their support and guidance. First and foremost, I would like thank Edward Anderson who has been a truly supportive advisor. He always had more confidence in me than I had for myself. I am very grateful for his advice not only on my research but also on academia and on pretty much everything else that matters. I had applied for UT to work with him and I am glad that I did.

Annabelle Feng has been a great mentor to me. She is an extraordinary researcher and I learned a lot from her. In those long ours we worked together, she was always friendly, helpful and patient. I am deeply grateful to her for her guidance.

When I decided to ask Jim Dyer's help about a technical issue in my second year paper, I was a bit nervous about approaching this highly senior faculty whom I barely knew. To my surprise, he turned out to be a very approachable and modest person with a great sense of humor. It was always a pleasure having discussions with him. I am grateful to him for his advice and guidance.

I feel very fortunate to have worked with Geoffrey Parker. He is not only a great researcher who always brings a fresh perspective to the discussion but also a mentor who supported me like an advisor throughout my studies.

Sridhar Seshadri has always inspired me with his enthusiasm and the breadth of his knowledge. I am glad that I had a chance to work with him and am thankful for his advice on my research projects.

I would also like to thank John Butler for his help on the first essay in this dissertation as well as for his valuable advice on my job talk.

My transition from a confused undergraduate to a PhD student targeting a career in academia has mostly to do with Yaman Barlas. He is an inspiring teacher and researcher who, without knowing, tricked me into academia. I am indebted to him for his guidance throughout my graduate studies.

I want to express my deepest gratitude to my parents and my sister for their never ending support. They have always been affectionate and understanding -even when I had to get work done during the limited time I visit them in summers. Without their trust in me, I would not have been able to do the least I have done.

Writing a dissertation was quite frustrating at times. My husband has been my biggest support while going through those phases. He was not only the most patient person in the world but also was great at reminding me the big picture and what matters the most. He has been my anchor and I cannot begin to tell how grateful I am for that.

Fortunately PhD life was not all about writing a dissertation in a poorly lit cubicle. I have met great fellow students whose support was indispensable and whose company made this journey a much more enjoyable one. I would like to especially thank Irem, Sree, Arshia, Fehmi, Saurabh and Josh for their support and friendship. Finally, I would like to thank those wonderful people at the Department Office who made our lives easy, especially Hillary Patterson for being so helpful in every step of the process.

Risk Mitigation Strategies for Project Management, Platform Development and Supply Chain Design

Publication No. ____________________

Burcu Tan, Ph.D.
The University of Texas at Austin, 2010

Supervisor: Edward G. Anderson Jr.

This dissertation studies strategies to mitigate the risks associated with operational and strategic decisions of a firm, particularly focusing on project management, product development and procurement decisions. In the first essay we develop two simulation-based methods to evaluate risky capital investment projects that involve managerial flexibility. Many risky projects are characterized by significant demand and operational risks (such as learning curve uncertainty) that are difficult to capture by simple stochastic processes. We propose using system dynamics simulations to estimate the cash flow resulting from these projects and build upon prior work on real options valuation in the decision analysis literature to develop two valuation algorithms. In the second essay we explore the technology investment decisions for platforms in markets that exhibit cross-network effects. We focus on the trade-off firms must make between investing new product development resources to increase a platform's core performance and functionality versus investments designed

to leverage the platform's cross-network effects. Abstracting from examples drawn from multiple industries, we use a strategic model to gain intuition about how to make such trade-off decisions under competition. In the third essay, we analyze the optimal procurement strategy of a firm that faces supply and demand risk. In particular, the firm can source from two unreliable suppliers with different delivery characteristics. We study the optimal order allocation policy shaped by the trade-offs between delivery leadtime, reliability and procurement cost. Further, we discuss the value of leadtime flexibility in supply risk mitigation and highlight the role of an inferior supplier in a firm's multi-sourcing strategy. The main contribution of this dissertation to the operations management literature is two-fold. First, it illustrates the role of effective risk mitigation through operational strategies of leadtime flexibility and supply diversification as well as through recognizing managerial flexibility. Second, it highlights the importance of leveraging third-party content development while making technology investment decisions for platforms in two-sided markets.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Risk mitigation is inherently a vital feature of operations management. With rapidly changing global markets and the increasing complexity of supply chains, effective mitigation of operational risks has gained further significance. In this dissertation, we consider some non-traditional methods of risk mitigation including purchasing real options to expand capacity, the use of diversified supply portfolios and the use of platform design to secure sufficient third party development. In particular, we study risk mitigation through operational strategies of leadtime flexibility and supply diversification as well as through recognizing managerial flexibility. Further, we examine the role of product development strategy in "getting both sides on board" in a two sided market.

In the first essay we develop two simulation-based methods to evaluate capital investment projects that involve managerial flexibility. Many important risky projects are characterized by significant demand and operational risks (e.g. learning curve uncertainty) embedded in non-linear feedback structures. Simple stochastic processes that are often used to value financial options may not capture the complex real-world behavior of these uncertainties.

We propose using system dynamics (SD) methodology to model the project uncertainty in order to take advantage of this methodology's capabilities in modeling complex feedback systems. The goal is improve the accuracy of the ultimate valuation by increasing the realism of the project model. Specifically, we use system dynamics simulations to estimate the cash flow resulting from risky projects and build upon prior work on real options valuation in the decision analysis literature to develop two valuation algorithms. We illustrate these algorithms with a model drawn from the wind power industry, which is characterized by numerous uncertainties and high managerial flexibility.

The first algorithm we propose, *the SD-based decision tree approach*, is based on transforming the system dynamics model of the project into an approximate decision tree that accounts for the managerial flexibility and then evaluating the tree using the risk adjusted discount rate. The second one, *the diffusion approximation approach*, involves approximating the cash flow uncertainty generated by the system dynamics simulations with a binomial decision tree and then evaluating the tree using risk-neutral valuation. We discuss the differences between these two approaches and describe under what conditions each one might be a more appropriate choice.

In the second essay, we examine the development of platforms in two-sided markets characterized by strong cross-network effects. Two-sided markets consist of one or more platforms that facilitate interactions between distinct groups of users. Examples include some of the most important industries in the economy such as video-game platforms, credit cards, cell phones, auction sites and operating systems. The distinguishing characteristic of two-sided markets is that the number of users on one side of the market depends on the number of users on the other side. For example, video-game developers will develop games for a platform only if the platform has a sufficiently large installed base. Likewise, all else being equal, gamers will prefer platforms with a greater variety of games. The presence of these "cross-network effects" often requires different strategies than those developed for traditional products and services.

In many cases, manufacturers of platforms in these two-sided markets face a trade-off between developing a high performance platform that matches the end-user's preferences and sacrificing some of those preferences in exchange for improved or less costly third party development capabilities. In particular, high performance is a competitive advantage in attracting the end-users however an aggressive investment in platform technology may make it more costly to develop content for the platform. This bears the risk of insufficient content development by third party developers, which in turn reduces end-users' willingness to pay for the platform. In this essay, we use a strategic model to gain intuition about how to make such trade-off decisions under competition.

Specifically, we study performance investment decisions of two competing platforms. We assume that higher performance makes third party content development more costly, either directly as in the case of video game platforms or indirectly by taking away resources that could have been used on improving third party development capabilities. We consider two different games based on the order of market entry: a simultaneous-move game where platforms enter the market at the same time and a sequential-move game where there is a leader and a follower in the market. Our results suggest that contrary to the conventional wisdom about "winner-take-all" markets, heavily investing in the core performance of a platform does not always yield a competitive edge when there are strong cross-network effects. In particular, offering a platform with lower performance but greater availability of content may be a better strategy.

In the third essay, we study dual-sourcing from unreliable suppliers with different delivery characteristics. Many firms resort to multiple-sourcing for various reasons. First of all, maintaining a diverse portfolio of suppliers allows for risk sharing. Due to the increasing complexity of supply chains, modern firms are much more vulnerable to disruptions in supply. The recession has exacerbated these concerns by increasing the capacity risk as suppliers take capacity offline by shuttering facilities, cutting shifts, and curtailing new production until the market returns. Consequently, supply diversification has become a common strategy to hedge against these possible shortfalls of supply. Second, multiple sourcing may provide leadtime flexibility. In other words, firms may make the best cost–responsiveness balance by keeping a cheap offshore supplier and an expensive onshore supplier. To fully take advantage of these benefits of multiple sourcing, it is critical to take into account the key dimensions of supply characteristics, mainly cost, reliability and leadtime. In this essay, we attempt to understand the joint effect of these dimensions on a firm's multiple-sourcing strategy.

In particular, we consider a firm that has two procurement sources, a fast supplier of one-period delivery leadtime and a slow supplier of two-period delivery leadtime. The capacity of each supplier is uncertain upon ordering which introduces a supply risk. We use dynamic programming to explore the optimal procurement policy for different levels of net inventory, cost parameters, and supplier reliability. We analyze how the firm allocates its total order between the two suppliers to mitigate supply and demand risks while facing trade-offs between leadtime, reliability and procurement cost. We shed light on the unique risk mitigation benefits associated with the slow supplier. Our results highlight the importance of incorporating cost, leadtime and reliability in a unified framework to evaluate supplier selection strategies.

Overall, this dissertation contributes to the existing operations management literature in two ways. First, it illustrates the role of effective risk mitigation through integrating leadtime flexibility and supply diversification as well as through recognizing managerial flexibility. Second, it highlights the importance of leveraging third-party content development while making technology investment decisions for platforms in two-sided markets. In Chapter 5, we present a summary of our results and discuss directions for future research.

Chapter 2

Evaluating System Dynamics Models of Risky Projects Using Decision Trees and Real Options Theory

2.1 Introduction

System dynamics (SD), beginning with the work of Forrester (1961), has been used to support project management for decades (Abdel-Hamid and Madnick 1991, Abdel-Hamid 1984, Pugh 1984, Roberts 1974, Lyneis and Ford 2007). Yet, the number of studies that use SD models to capture the value of managerial flexibility within risky projects remains small (Lyneis and Ford 2007). The real options valuation approach is the state-of-the-art method to value capital investment projects. It applies financial options theory to value options derived from managerial flexibility, which are called "real options" to reflect their association with real assets rather than with financial assets (Myers 1987, Trigeorgis and Mason 1987, Trigeorgis 1991, Dixit and Pindyck 2001). Accounting for the value of managerial flexibility such as the right, but not the obligation, to expand, postpone, or terminate a project during its implementation is essential to provide a good estimate of the value of risky projects. Hence, integrating SD models of projects into the existing real-options valuation methods may offer many potential benefits.

The main benefit of using SD models for the valuation of risky projects is the increased realism of the project model itself. Traditional methods to value managerial flexibility model project uncertainty by assuming an analytically tractable stochastic process (e.g. McDonald and Siegel 1986, Paddock, Siegel and Smith 1988, Pindyck 1991, Capozza and Li 1994, Trigeorgis 1991, Kogut and Kulatilaka 1994). Yet, these simple stochastic processes may not capture the complex real-world behavior of uncertainties that result from non-linear feedback structures, such as rework and learning curves. Modeling the structure that produces this complex behavior with SD methodology may improve the accuracy of the ultimate valuation. Another advantage of using SD methodology is its flexibility in defining complex feedback systems and separate stochastic effects, which is quite beneficial in dealing with multiple and potentially interacting sources of uncertainty. In addition, describing the distribution of uncertainty around SD variables is intuitive thanks to the methodology's emphasis upon the use of concrete variables that correspond to "real" phenomena (Sterman 2000). As a result, SD provides clearer insights into the drivers of the effect of a strategic action (Johnson, Taylor and Ford 2006).

There is a burgeoning literature that recognizes these benefits. Ford and Sobek (2005) build a product development project model that uses real options concepts to manage product design risk. Ford and Bhargav (2006) examine the relationship between project management quality and the value of flexible strategies. Johnson et al. (2006) use an SD model to value flexibility in a large petrochemical project. These papers provide examples of how managerial flexibility can be incorporated into SD models of projects by analyzing projects with a single option. They do not focus, however, on formalizing an algorithm to estimate the value of projects in which managerial flexibility is characterized by complex option structures.

The difficulty in evaluating complex option structures within SD models lies in the necessity to optimize a sequential decision process, which typically involves backwards induction (Bertsekas 2005), a technique incompatible with most SD simulation evaluation algorithms. In contrast to system dynamics, decision tree analysis provides an intuitive approach commonly used to model sequential decision processes (Clemen 1997) and is compatible with backwards induction. The decision tree models are also appealing because they are easy to explain to non-practitioners.

To take advantage of the complementary strengths of system dynamics and decision analysis in representing stochastic models and decision processes respectively, we propose two decision-tree based algorithms, *SD-based decision tree approach* and the *diffusion approximation approach*, to evaluate SD models of projects that involve managerial flexibility. The algorithms we propose are based upon first formulating a system dynamics model of the project and then transforming the cash flow data obtained from the model into a decision tree. By leveraging the decision analysis literature (Clemen 1997), the algorithms enable a backwards induction (Bertsekas 2005) solution approach to evaluate the project. This approach is inherently attractive because Nau and McCardle (1991) and Smith and Nau (Smith and Nau 1995) demonstrate that decision tree analysis can duplicate real options based valuations under certain conditions.

The SD-based decision tree algorithm's transformation is similar to the method described in Osgood and Kaufman (2003) and Osgood (2005)-which is to the best of our knowledge, the only explicit use of multiple-decision decision trees in the SD literature-with the important distinction that the proposed new algorithm incorporates chance events dictated by endogenous dynamic processes as opposed to exogenous scenarios generated by strictly exogenous processes (the distinction will be clarified in Section 2.3). This enables the SD-based decision tree algorithm to handle multiple endogenous sources of uncertainty efficiently with less vulnerability to "the curse of dimensionality". The resulting decision tree is then evaluated using a risk-adjusted discount rate for the project, which in practice is often the company's weighted average cost of capital (WACC).

One caveat with the SD-based decision tree approach is that the optimization that occurs at the decision nodes changes the expected future cash flows and thereby alters the risk characteristics of the project so that the risk-adjusted discount rate for the project without options may not be appropriate after the real options have been included in the model. To remedy that, the second algorithm we propose, *the diffusion approximation approach*, modifies the SD-based decision tree approach by using concepts from Copeland and Antikarov (2001) and Brandao et al. (2005). The diffusion approximation approach avoids the problem of selecting an appropriate risk-adjusted discount rate for the analysis by using a "risk neutral" valuation that provides a more accurate estimate of the market value of the project.

The remainder of the paper is as follows. Section 2.2 provides a motivating example for this paper, an alternative energy windmill project, and describes an SD model for that project. This SD model includes many features that would be prohibitive to represent in real options solution approaches based on closed-form stochastic calculus or on spreadsheet models. Section 2.3 presents the SD-based decision tree algorithm by illustrating the transformation of SD simulation data into a decision tree that incorporates the project's "real options". Section 2.4 introduces the diffusion approximation algorithm. Section 2.5 illustrates the steps of valuing the example project using the diffusion approximation algorithm. A comparison of the two valuation approaches as well as a discussion of some limitations is provided in Section 2.6, followed by a short conclusion in Section 2.7.

2.2 A Motivating Project and its System Dynamics Model

Concern about the high price and the environmental impacts of fossil fuel use has increased support for alternative energy technologies (AETs) such as wind power. However, developing these technologies has proven problematic, in part because of the difficulty in estimating the return from investing in such projects. AETs compete in the electricity generation market, the price dynamics of which have a correlation of 0.85 with natural gas prices (EIA 2007), and in which the dominant, conventional technology is the combined-cycle natural gas plant. Hence, the price of natural gas is a major determinant of how competitive AETs will be. Yet, natural gas prices are uncertain and cyclic and are influenced by geopolitical and macroeconomic short-term factors.

The cost of developing a new technology is also uncertain. The bulk of most AETs' cost structures lies in their non-recurring costs. As for many other new technologies (Argote 1999), these costs are typically reduced significantly with each doubling of the cumulative capacity installed. However, the steepness of this "learning curve" and its final "plateau" are generally unknown ex ante. For example, estimates for wind power learning curves vary from 12% to 23% cost reduction per doubling of the cumulative capacity installed (Junginger, Faaij and Turkenberg 2005, Musial and Butterfield 2004).

We will use a hypothetical wind power project to illustrate an SD-based valuation method that handles multiple uncertainties in a reliable and realistic manner. The SD methodology has been used to develop models that can plausibly track both fossil fuel prices (Morecroft and van der Heijden 1992, Naill 1992, Davidsen, Sterman and Richardson 1990, Sterman and Richardson 1985, Sterman 1981, Ford 1997) and technology "learning curves" (Anderson and Parker 2002) as well as other forms of project risk embedded within endogenous nonlinear feedback loops. The complex interactions among these uncertainties are not captured easily with stylized models (Forrester 1975). In principle, the model for this project, like any other system dynamics model, might also be evaluated by a spreadsheet. However, the number of interdependent dynamic equations involved may make this impractical for any non-trivial model. Furthermore, such a spreadsheet model would conceal the structure and therefore be potentially error-prone.

The illustrative example in this paper is based on a hypothetical firm's effort to evaluate an investment opportunity to build a 40-MW wind farm with the option to add 50 MW within the first 4 years. The firm can also delay the beginning of the project by up to 2 years. The expansion option may be considered after the 40-MW wind farm comes online, which takes a year. The option can be distributed over the remaining 2 years, but due to economies of scale, the firm does not want to have less than 25 MW built at a time. So, the firm can either expand high (build all 50 MW at once) or expand low (build 25 MW one year) with the option to build another 25 MW in the successive year or suspend investment, i.e. continue operating the wind farm at its current capacity. The details of the project specifications are provided in Table 2.1[1].

Project Specifications	
Capital Costs ($/kW)	$1,015
Fixed Operation and Maintenance Costs ($/kW/yr)	$20
Variable Operation and Maintenance Costs ($/MWh)	$1
Capacity Factor (%)	30%
Shaping and Integration Costs ($/MWh)	$7
Project life	20 years

Table 2.1: Project Specifications

This example is kept simple for the sake of clarity in exposition, but it

[1]Cost figures are adapted from Northwest Power and Conservation Council (2007)

contains sufficient embedded stochastic processes that make system dynamics models attractive. The proposed algorithm can handle this as well as much more complex models (albeit within certain limitations discussed in Section 2.6).

Figure 2.1 is a sector diagram of a generic SD model for evaluating a wind-power AET project[2]. Implicitly, there is a two-tier supply-chain structure in the model: The supplier, who installs the equipment for the energy plant (e.g. windmills), and the generating firm, who evaluates the investment opportunity. As the firm acquires more capacity, the supplier "learns-by-doing" which leads to increased productivity and lower installation costs. Hence, the more the generating firm invests, the lower the costs it will face in its future investments. Similar feedback relations in this context have been modeled in Ford (2006) and Vogstad (2004). The supplier also takes some advantage of global technological improvements, which is approximated as the impact of a "global" capacity acquisition level.

The model is built for a medium scale firm, which is a price-taker. As described earlier, electricity price dynamics are correlated with the dynamics of natural gas price. The gas price sector is based on the feedback loops described in Sterman et al (1988). State and federal regulations also play a key role in determining the profitability of an AET investment project. For example, renewable energy producers currently receive a 1.9 cent benefit per

[2] A description of the model can be found in Appendix A

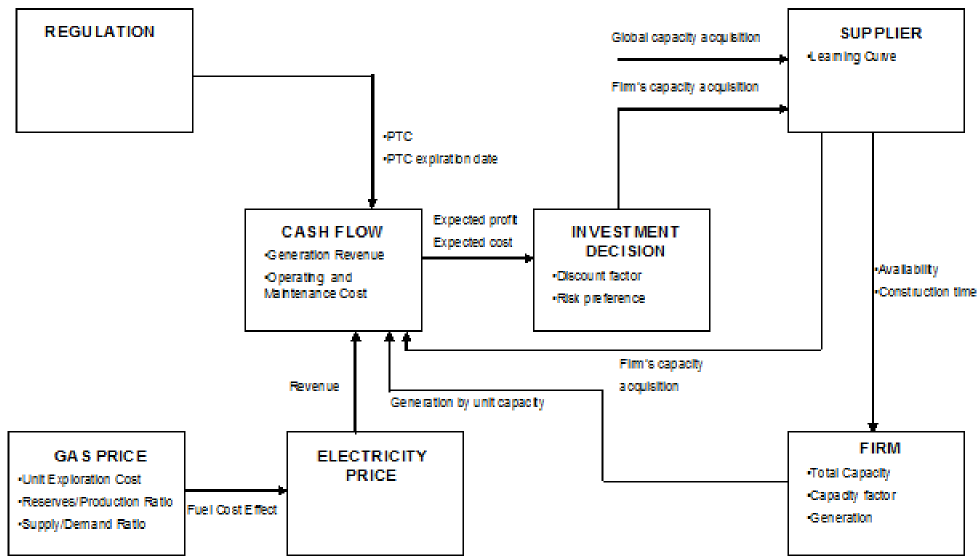


Figure 2.1: Sector Diagram of the SD Model

kilowatt-hour of generation, known as the production tax credit (PTC). The uncertain expiration date of the PTC has been a major consideration in wind capacity investment decisions.

Three major uncertainties are captured by the model: natural gas price, the learning curve and the expiration date of the production tax credit (see Table 2.2). Natural gas price uncertainty has several components. The *demand growth ratio* is the rate that "base demand" for natural gas is assumed to grow. The *base demand* is assumed to grow continuously. It is primarily a function of population growth less any potential reduction in energy usage intensity. *Initial undiscovered resources of natural gas* is another major uncertain variable that determines the gas supply and hence the future gas prices. In addition to these factors, the natural gas supply is exposed to random disruptions whose frequency, size and duration are uncertain.

Uncertain Parameters	
Gas Price Model Parameters	**Units**
Demand Growth Ratio	1/Month
Initial Undiscovered Resources	tcf
Supply Disruption Size	1/Month
Supply Disruption Frequency	Dimensionless
Average Length of Supply Disruptions	Month
Learning Curve Parameters	**Units**
Steepness of the Learning Curve	Dimensionless
Weight of Global Learning	Dimensionless
Official Expiration Date of PTC	Month

Table 2.2: Uncertainties Captured in the Model

The learning curve uncertainty has two components: The *steepness of the learning curve* determines how fast the cost of capacity drops with doubling of the cumulative capacity installed. The *weight of global learning* determines how much the supplier benefits from the technological improvements (global capacity acquisitions) elsewhere. Other uncertainties (e.g. wind speed, the minimum possible capacity cost, etc.) could easily be incorporated without increasing the complexity of the solution procedure.

2.3 SD-based Decision Tree Approach

In this section, we describe the SD-based decision tree approach. The first step of the algorithm is identifying the "real options" in the project.

At each period, the manager needs to decide whether or not to exercise the available options; hence there is a sequence of decisions to be made throughout the horizon of the project. Each possible sequence of these decisions is called a *decision sequence*. In the example problem, the firm has several managerial options that are valid at different time periods: invest (I), delay (D), expand high (H), expand low (L) and suspend further investment (S). The time period is set at one year, so $t = 1$ represents the end of the first project year. Hence, these options result in the 10 decision sequences that can be seen in Table 2.4.

Steps of the Algorithm	
Step 1	*Identify the managerial decision sequence*
Step 2	*Build the deterministic SD model that captures the project dynamics*
Step 3	*Model the uncertainty by specifying the random variables and their distribution*
Step 4	*Run Monte Carlo simulations of the SD model for each decision sequence*
Step 5	*Obtain the discrete distribution approximations for the first period cash flow distribution for each decision sequence*
Step 6	*Obtain the conditional discrete approximations for the remaining periods for each decision sequence*
Step 7	*Solve the decision tree by backwards induction using the risk adjusted discount rate*

Table 2.3: Steps of the SD-based decision tree algorithm

The second step is building the deterministic SD model that captures the project dynamics. In the example problem, the model must capture the AET capacity investment, fossil fuel prices, and learning curve dynamics.

Then, as Step 3, the underlying uncertainty of the project has to be modeled by specifying the random processes and their distributions. In this case, there are multiple sources of uncertainty as presented earlier in Table 2.2.

Decision Sequence ID	**Decision in 2008**	**Decision in 2009**	**Decision in 2010**	**Decision in 2011**
1	I	N/A	S	S
2	I	N/A	S	L
3	I	N/A	S	H
4	I	N/A	H	S
5	I	N/A	L	S
6	I	N/A	L	L
7	D	I	N/A	L
8	D	I	N/A	H
9	D	I	N/A	S
10	D	D	I	N/A

Table 2.4: Decision Sequences for the Example Problem

Next, the valuation problem is translated into a decision tree. Basic components of a decision tree are as follows: Square nodes are the *decision nodes*, which represent the decisions to be made at a particular time, like "invest high or suspend". Branches leaving a decision node represent the decision alternatives. Circular nodes are the *chance nodes*, which represent the uncertainties underlying the project. Branches leaving a chance node represent possible outcomes of an uncertain event, and we use the SD-based decision tree approach that requires any continuous uncertainty to be approximated by a discrete probability distribution[3]. Triangular nodes are the *terminal nodes* that depict the final outcome of a particular scenario after all decisions have been made, all uncertainty has been resolved and all payoffs are received.

The fourth step of the algorithm is running Monte Carlo simulations for each decision sequence in Table 2.4 (It is straightforward to impose these decision sequences in the SD model with a few additional if-then-else type equations). A Monte Carlo run for a specific decision sequence gives a cash flow distribution for each period. In the example problem, the time horizon is 20 periods; hence, we obtain 20 cash flow distributions for each decision sequence in Table 2.4, making a total of 200 distributions. For example, the distribution displayed in Figure 2.2 is the distribution of the 7th period cash flows obtained under the decision sequence "invest-suspend-suspend" after 1000 iterations of the simulation model.

In traditional decision analysis models, uncertainty is modeled by assigning probabilities to each branch leaving a chance node. These *exogenously* assigned probabilities may represent subjective beliefs about the likelihood of a specific "event" represented by the chance node (e.g. the probability of PTC being suspended) or they may be risk-neutral-probability measures[4] of price

[3]Modern decision tree software does allow a chance node to be represented by a continuous probability distribution, and can use Monte Carlo simulation to obtain a solution for this case. However, this approach becomes computationally challenging for complex decision trees with multiple chance nodes, and it also requires the calculation of the ending node outcome for each iteration of the simulation. Since we use an SD model to calculate the ending node outcome, this could require running the SD model thousands of times.

[4]Risk-neutral measure is an important concept in the context of mathematical finance and risk neutral valuation is an important general principle in option pricing (Hull 2006).

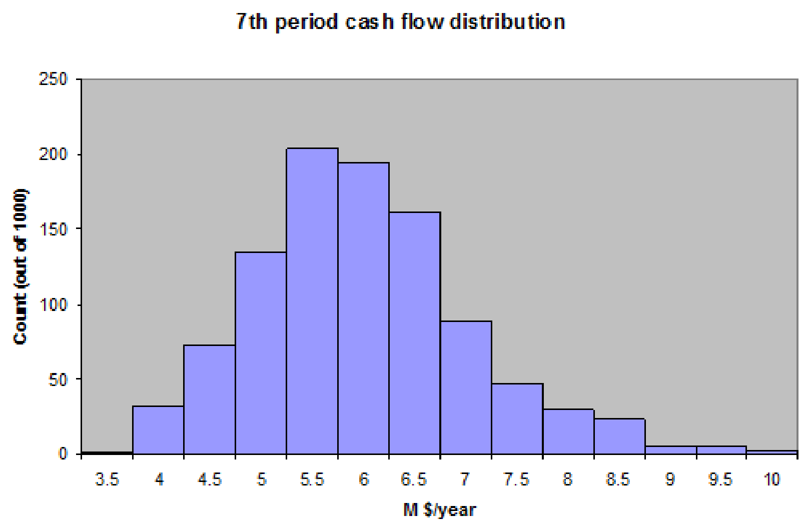


Figure 2.2: The Distribution of the 7th Period Cash Flows under the "I-S-S" Decision Sequence

uncertainties derived from market data. However, in our approach exogenous scenarios such as PTC suspension are incorporated at the stage of building the stochastic SD model. Hence, instead of building a decision tree that explicitly models every source of uncertainty as separate chance nodes, we take a different approach. In our method, a chance node represents the distribution of the random cash flow accumulated during a given period t, which is obtained from the Monte Carlo simulation of the SD model for the corresponding decision sequence. These cash flow distributions are the outcome of the random

A risk-neutral measure is a probability measure in which today's arbitrage-free price of a derivative security is equal to the discounted expected value (under the measure) of the future payoff of the derivative. The measure is in general different than the "physical" measure of probability and is employed to determine the worth of derivative securities. Please refer to Dixit and Pyndick (2001) for further details.

processes that are shaped in the SD model.

This representation of chance nodes has two advantages. First, it allows multiple uncertainties in a time period to be represented by a single uncertain variable (the cash flow in that time period) without changing the size of the tree and without adding any computational burden. The only cost of adding more uncertainties is the increased run time for the Monte Carlo simulations. Second, the approach simplifies handling valuation problems that involve path-dependent stochastic processes, such as the investment cost at period t which is a function of the cumulative capacity investment up to t. Such processes are represented effectively within an SD model. Since the cash flow distributions are obtained through SD simulations, the proposed approach is powerful in handling complex, path-dependent stochastic feedback structures.

The cash flow distributions obtained from the Monte Carlo runs are continuous. In order to place them into the tree as chance nodes, one needs to transform them into discrete distribution approximations without losing too much precision. We use the bracket median approximation technique (Clemen 1997) to obtain a k-point discrete distribution approximation: First, the distribution is divided into k equally likely intervals. Then the median of each interval is determined. Hence, the continuous distribution is approximated with k median values that are each equally likely to occur (with probability $1/k$). The choice of k is a critical one. If too few intervals are chosen, the conformity of the discrete cash flow distribution to the continuous distribution will be low, which may lead to the recommendation of an alternative that would not be optimal for the true cash flows. If too many are chosen, the computations involved will become prohibitive. The issue of the appropriate level of detail in a model is common to all applications, and the practitioner needs to resolve this trade-off on a case-by-case basis. As an example, when a project has a relatively short horizon or when managerial flexibility is limited to a certain phase of the project rather than being spread out in the entire horizon (such as an option to delay), dimensionality is less of a problem, and it is feasible to use a bigger k. Yet, if the analysis does not demand high fidelity to the tails of the distribution, a bigger k may not be necessary.

The fifth step of the algorithm is obtaining k-point discrete distribution approximations for the *first* period cash flow distributions of each decision sequence by using the bracket median method. In the example problem, three-point discrete approximations are used. Depending on the firm's decisions and on how the natural gas price, capacity cost and tax credit uncertainties evolve, the cash flow at the end of each period may be termed *high*, *medium* or *low* (each of these could possibly be less than zero). Note that the cash flow distributions for period t are conditional on the cash flow distributions of period $t-1$, as well as on what *decision sequence* is chosen. For example, the *high* level at period t given that the cash flow was low in period $t-1$ is, in general, different from the high level given that the cash flow in period $t-1$ was high.

In Step 6, the discrete conditional probability distributions are computed using the following procedure:

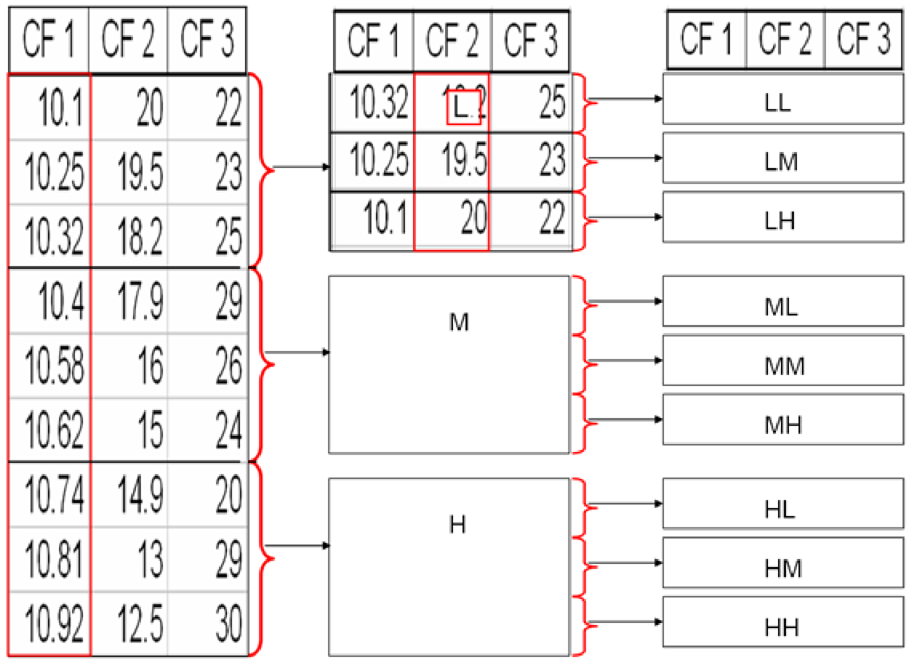


Figure 2.3: Illustration of the Conditional Distribution Computations for a Given Decision Sequence

Step 6.a : Divide the initial table of cash flows into k tables designed for the designated levels of the first period cash flows (i.e. *high*, *medium*, *low*). Make sure to preserve the sample path structure: If a realization of $\tilde{C}_1$ falls into the bracket *high* (as a result of Step 5), then carry that entire sample path $\tilde{C}_1, \tilde{C}_2, ..., \tilde{C}_T$ over the table for *high*. As a result of this step, k tables with N/k rows and T columns are obtained where N is the size of the Monte Carlo simulation and T is the number of periods (in this case, years) in the valuation problem.

Step 6.b : Apply the bracket median approximation to each table obtained in Step 6.a in order to discretize *the second period cash flows*. Then, further divide each table obtained in Step 6.a into k tables with N/k^2 rows and T columns with respect to second period cash flows $\tilde{C}_2$, obtaining k^2 tables (*high-high*, *high-medium*, *low-high*, etc). The process is illustrated in Figure 2.3.

Step 6.c : Repeat this procedure until the last period. If there are no decision nodes after a certain period t, one can apply the procedure to the present value of cumulative cash flows after t instead of carrying the procedure until the last period. This would result in losing the volatility information after t but would not change the optimum decision sequence and the expected net present value.

Note that these operations for obtaining the conditional probabilities must be carried out for each decision sequence. Fortunately, this can be accomplished with the help of a simple Visual Basic macro. The decision problem is represented with the decision tree shown in Figure 2.4 using the software program DPL. Note that this representation of the tree is compact. If a single chance or decision node follows multiple branches from a predecessor node (like the highlighted nodes in Figure 2.4), it indicates that the chance or decision node is actually appended at the end of each branch of the predecessor node. For example, the highlighted decision node "OPTION T2" is represented to follow the "Medium" branch of "Cash Flow T1" but actually it is appended at all the branches of "Cash Flow T1". The paths that are not explicitly depicted in this compact representation are hidden for visual clarity but are not excluded from the solution procedure.

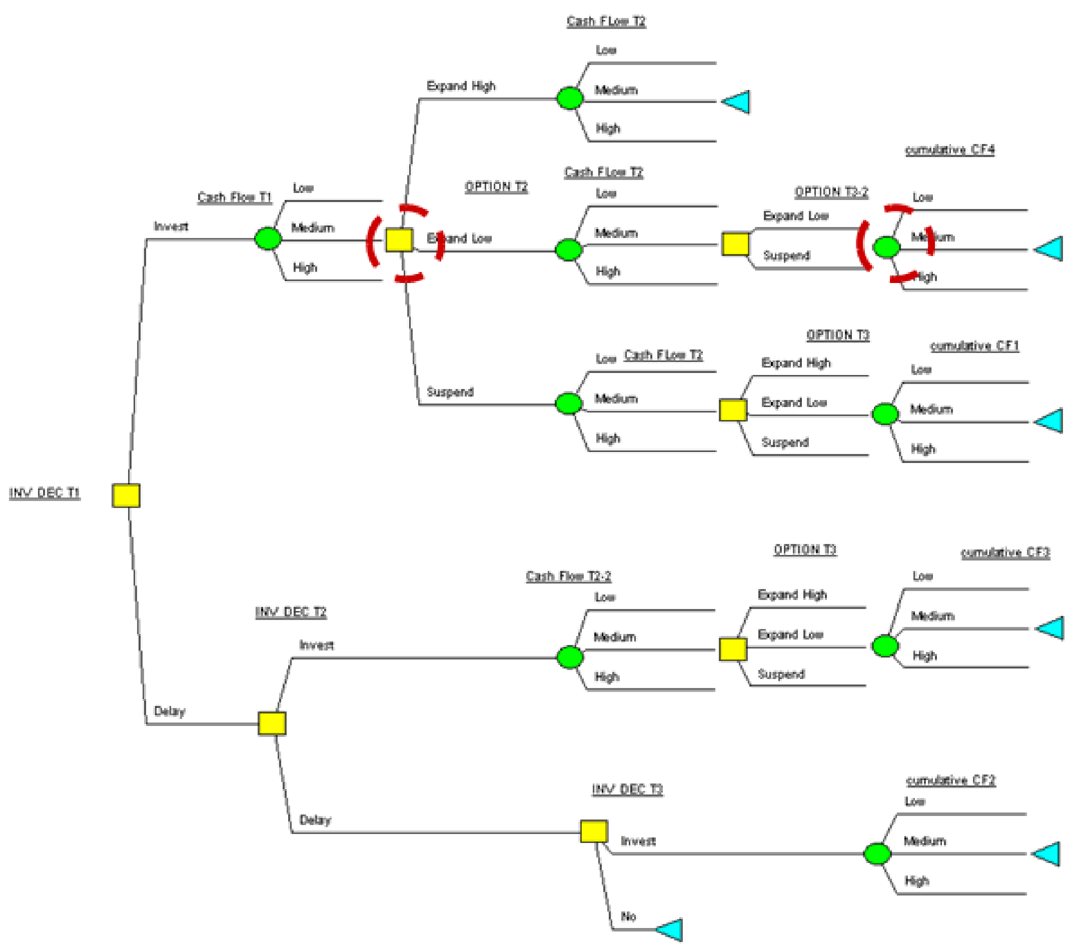


Figure 2.4: Decision Tree Representation in DPL

The last step of the algorithm is solving the decision tree. After obtaining the net present value (NPV) for all the terminal nodes of a decision sequence for each decision sequence, we evaluate the decisions at the final period T and then move backwards using backwards induction. In practice, this step is handled easily by using decision analysis software such as DPL.

The example problem is solved using the DPL decision analysis software with a discount rate equal to the WACC of the firm, which is assumed to be 10%. The optimal policy is highlighted in bold in Figure 5. The expected PV is \$55.386 million and the expected NPV of the project is \$14.786 million. The expected NPV of the project without options (which is the expected NPV of the strategy "invest-suspend-suspend") is \$7.79 million; hence the combined value of the options is approximately \$7 million. The positive NPV suggests that the project should be undertaken. If the cash flow in the first period is *low* or *high*, the firm should expand low in two consecutive periods taking full advantage of the learning curve effect. High cash flow corresponds to high natural gas prices and continuing production tax credit. If the cash flow is *medium*, expanding high is optimal; yet, expanding low in two consecutive periods yields practically the same NPV (\$12.81 vs. \$12.77 million).

The value of a specific option associated with each decision node may be determined simply by rebuilding the tree without that option and solving that simplified tree. An estimation of the expected value of that option is the difference between the value of the project with the option and without the option. If the option has a cost, the firm should acquire the option if this estimated value is higher than the cost of the option.

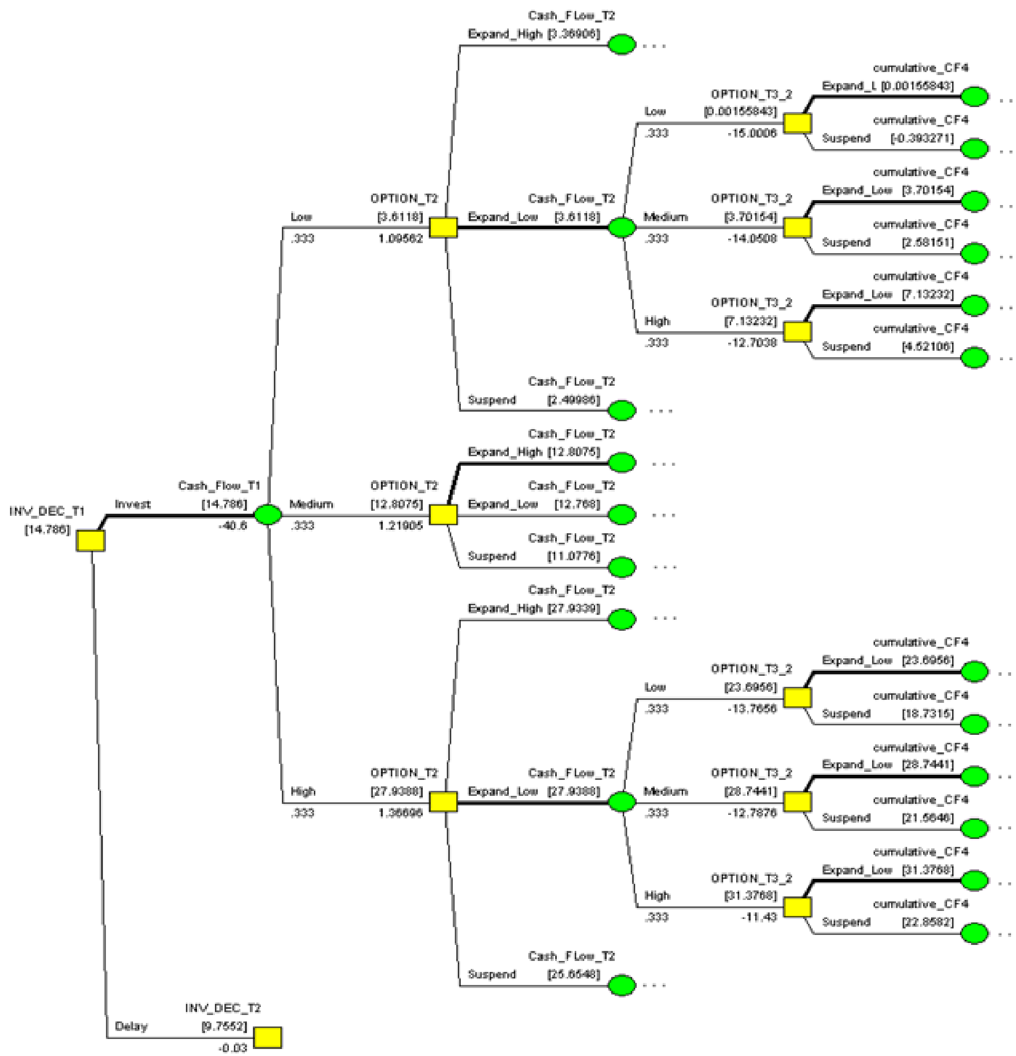


Figure 2.5: The Optimum Policy for the Example Problem

2.4 The Diffusion Approximation Approach

One major limitation of the SD-based decision tree approach is that the risk-adjusted discount rate for the project without options is used as the discount factor for the entire decision tree (Teisberg 1995). Essentially, the risk-adjusted discount rate that a financial analyst should choose to value the project with options may be different from the one he should choose to value the project without options because of the alternatives' different risk levels. An analysis based on a single discount rate may be unreliable especially if the valuation results are highly sensitive to the discount rate.

The classical approach to incorporating market information on risk is based upon identifying a replicating portfolio for the project under consideration and using the volatility information of this replicating portfolio to obtain an appropriate discount rate for the project. However, the replicating portfolio assumption is difficult to use in practice when evaluating individual corporate investments because it is hard to find a single replicating asset or even a portfolio of publicly traded assets with returns that are perfectly correlated with those from the project (Borison 2005), and the appropriate replicating portfolio may change if the risk of the project is changed by the addition of options.

This criticism can be overcome by using a decision-tree based real options valuation method developed by Copeland and Antikarov (2001) and modified by Brandao, Dyer and Hahn (2005). The method is a decision tree based approach that combines any number of uncertainties into one combined stochastic process representing the project value and it avoids the need to find a replicating portfolio. We will make some modifications to this method and call it the diffusion approximation approach. Copeland and Antikarov (Copeland and Antikarov 2001) and Brandao, Dyer and Hahn (2005) suggest the use of a Monte Carlo simulation of a pro forma spreadsheet model to obtain an estimate of the risk associated with the project without options, which is then used to construct the required decision tree. Instead we substitute simulation runs from an SD model and obtain a reliable, and theoretically correct (from the point of view of the finance literature) valuation of the investment projects.

In its simplest form, the diffusion approximation method assumes that the changes in the project's value over time approximately follows a geometric Brownian motion (GBM) diffusion process, which is a standard assumption in the finance literature (e.g. Copeland and Antikarov 2001). The method relies on the *market asset disclaimer* (MAD) assumption[5], which assumes that the value of the project without options is the best unbiased estimator of the market value of the project. Hence, the present value of the project without options is taken as the market price of the project as if it were traded (Copeland and Antikarov 2001). Then, the value of the project without options is assumed to change over time according to a GBM process, which is the same process used to model the changes in the price of a stock when the Black and Scholes (1973) option pricing model is used. The assumptions behind the GBM model may not hold for all projects, in which case other models

[5]The MAD assumption is used in order to create a complete market for an asset that is not traded in the market. It is a strong modeling assumption made to justify the use of risk-neutral valuation. Nevertheless, it eliminates the reliance on the existence of a replicating portfolio. Instead, it uses the project itself as the twin security and is claimed to "make assumptions no stronger than those used to estimate the project NPV in the first place" (Copeland and Antikarov 2001, p. 67). For further discussion of the MAD assumption, see also Borison (2005) and Smith (2005).

of stochastic processes (e.g. mean reverting) may be used (Brandao, Dyer and Hahn 2005, Hahn and Dyer 2008, Wang and Dyer 2009).

We use the GBM model because the use of a binomial lattice approximation to a GBM process is well established in the literature (Hull 2006) and it is straightforward to build a corresponding decision tree to value the project once the parameters of the process are provided. In the decision tree representation, the project values are discounted with the risk-free rate since the risk-neutral probabilities are used, which eliminates the need to estimate different risk-adjusted discount rates as options are added to the project.

The first three steps of the diffusion approximation approach are similar to the SD-based decision tree approach: 1) Identify the decision variables 2) Build the deterministic SD model 3) Specify the distributions of the uncertain variables. The next three steps involve calculating the parameters of a GBM approximation of the project uncertainty. The critical parameters required to model this approximation are the present value (PV) of the project without options, cash flow payout rate in each period t, δ_t, volatility of the project returns, σ, and the risk free rate, r.

Step 4 of the diffusion approximation algorithm is calculating the expected PV of the project without options by using a DCF analysis. To do that, we run Monte Carlo simulations of the project without options and obtain the PV for each iteration (sample path) using a risk-adjusted discount rate μ , such as the weighted average cost of capital (WACC) [6] of the firm. Then, the average of these iterations is used as an estimate of the PV of the project.

Step 5 of the diffusion approximation approach is obtaining the cash flow payout rate δ_t in each period. The cash flow payout rate is the ratio of the cash flow in a given period t to the value of the project in that period. The project value at time t is simply the present value of the *remaining* project cash flows. Note that each iteration j of the Monte Carlo simulation gives a sample path of cash flow obtained in each period t, denoted by $C_{t,j}$. Let $\tilde{V}_t$ and $\tilde{C}_t$ be random variables representing the remaining project value and the cash flow in period t, and $\bar{V}_t$ and $\bar{C}_t$ be their corresponding means. For each iteration j of the Monte Carlo simulation, the project value $V_{t,j}$ in each time period t is given by:

$$V_{t,j} = \sum_{i=t}^{T} \frac{C_{i,j}}{(1+\mu)^{i-t}} \tag{2.1}$$

where μ is the risk adjusted discount rate and T is the number of periods. Accordingly, the cash flow payout rate in period t is defined as:

$$\delta_t = \frac{\bar{C}_t}{\bar{V}_t} \tag{2.2}$$

In the algorithm, the cash flow payout rates are used to calculate the cash flows that are paid out at the end of each time period as a function of the project value. In line with the previous work, we assume that the cash flows vary over time reflecting the uncertainty in the project value, but that in each time

[6]See Brandao et al (2005) for a discussion on the choice of the discount rate for this step.

period they are a constant fraction (δ_t) of the remaining value of the project (Copeland and Antikarov 2001, Brandao and Dyer 2005, Brandao et al. 2005).

The sixth step of the algorithm is estimating the volatility (σ) of the project returns. Smith (2005) suggests an approach to estimating this volatility that can be used in the SD simulation environment. First, we model the GBM approximation of the project value setting $V_0 = PV$ and using:

$$dV_t = (\mu V dt + \sigma V dz)(1 - \delta_t) \tag{2.3}$$

where $dz = \epsilon\sqrt{dt}$ and $\epsilon \sim N(0, 1)$. Once the change in the uncertain project value is defined, we can easily model the random cash flow $\tilde{C}_t$ by using the assumption $\tilde{C}_t = \tilde{V}_t * \delta_t$. Then, we search for a volatility that best mimics the uncertainty in the original SD model; i.e., that minimizes the difference between the cash flow distributions generated by the original SD simulation model and the cash flow distributions obtained by the GBM approximation. One way of doing this is comparing the 10th, 50th and 90th percentiles of the cash flow distributions obtained through the GBM approximation *for a given volatility* to the corresponding values given by the original SD model. Note that this results in $3T$ pairs to be compared, where T is the number of periods. Sum of squared errors between these pairs provides a measure of fit for the GBM approximation under the given volatility. We repeat this procedure for a predetermined set of candidate values for volatility. Then, we choose the value that minimizes the sum of squared errors.

Once the volatility of the project returns is determined, we have all the parameters required to model the GBM that characterizes the project uncertainty. However, in order to represent this continuous stochastic process on a decision tree, a discrete approximation is needed. The binomial approximation to the GBM process serves this purpose by representing the GBM process with a binomial lattice. A binomial lattice is a probability tree with binary chance branches that go up (u) or down (d) with the unique feature that the outcome resulting from moving up and then down is the same as the outcome from moving down and then up (Figure 2.6). In particular, the binomial lattice model assumes that with probability p the value of the project V will go up to Vu, and with probability $1 - p$ it will go down to Vd at the end of one period. The parameter u is greater than 1 (reflecting a proportional increase), whereas d (reflecting a proportional decrease).

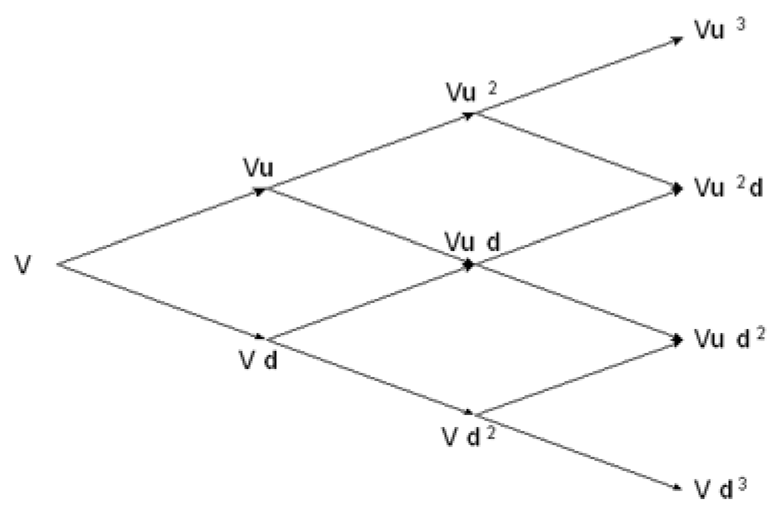


Figure 2.6: A Binomial Lattice

The next step of the algorithm involves calculating the parameters of the binomial lattice. The initial value V of the project is approximated by the PV of the project without options obtained in Step 4. In order to calculate the remaining three parameters u, d and p, it is sufficient to know the volatility σ of the GBM process, which is estimated in Step 6, and the risk-free discount rate r:

$$\begin{aligned} u &= e^{\sigma\sqrt{\Delta t}} \\ d &= 1/u \\ p &= \frac{1 + r\Delta t - d}{u - d} \end{aligned} \tag{2.4}$$

where Δt is the time period used in the binomial lattice. The probabilities p and $1 - p$ are the probabilities that a risk-neutral investor would assign to the two outcomes; therefore they are often called "risk-neutral" probabilities. Assigning risk-neutral probabilities enables using the risk-free rate to discount the cash flows; hence, we avoid estimating different risk-adjusted discount rates as options are added to the project.

The lattice may also be "unfolded" and represented as an equivalent binomial tree. This increases the number of endpoints in the model decreasing computational efficiency, but allows these problems to be solved using "off the shelf" decision tree software with a simple and intuitive visual representation. In this algorithm, we choose using the decision tree approach instead of a binomial lattice in order to take advantage of the aforementioned complementary advantages of decision tree analysis and system dynamics.

The eighth step of the algorithm is building the binomial tree representing the project without options, which requires calculating the project value and the cash flow at each period and each state up (u) and down (d). First, note that at each period a fraction δ_t of the project value is paid out as cash flow diminishing the value of the project to $V_t * (1 - \delta_t)$. Accordingly, the project value in the subsequent period is calculated as follows:

$$V_{t+1}^{u} = V_t(1 - \delta_t)u \tag{2.5}$$

$$V_{t+1}^{d} = V_t(1 - \delta_t)d \tag{2.6}$$

where u and d are given by (2.4). The cash flow $C_{t,k} = V_t^k * \delta_t$ at state k and period t can be discounted at the risk free rate r since risk-neutral probabilities are used. Hence, the *discounted* cash flow that is paid out at each time period t and at each state k is given by:

$$C_{t,k} = \frac{V_t^k \delta_t}{(1 + r)^t} \tag{2.7}$$

Once the project without options is modeled with a binomial decision tree (Step 8), options can be added in the form of decision nodes. For example, an abandon option can be modeled by adding a decision node without any subsequent chance nodes (i.e. no further cash flows), whereas simple expansion and contraction options can be modeled as percentage changes in the cash flows (for details, see Brandao et al. 2005).

These eight steps are summarized in Table 2.5.

Steps of the Algorithm	
Step 1	*Identify the managerial flexibilities*
Step 2	*Build the deterministic SD model that captures the project dynamics*
Step 3	*Model the uncertainty by specifying the random variables and their distributions*
Step 4	*Run Monte Carlo simulations of the SD model for the project without options and calculate the present value*
Step 5	*Obtain the cash flow payout rates for each period*
Step 6	*Estimate the volatility of the project returns*
Step 7	*Calculate the parameters of the binomial approximation to the GBM process (i.e. u, d and p)*
Step 8	*Build the binomial tree*
Step 9	*Add the options and solve the decision tree by using the risk free rate*

Table 2.5: Steps of the Diffusion Approximation Algorithm

2.5 Illustration of the Diffusion Approximation Algorithm

We will use the same hypothetical wind power project described in Section 2.2 to illustrate the diffusion approximation algorithm. The first three steps of the diffusion approximation algorithm are the same as explained in Section 2.3. At Step 4, the expected PV of the project without options at time $t = 0$ is calculated using a DCF analysis. To do that, we use the Monte Carlo simulation of the strategy "invest-suspend-suspend", which reflects the project without options. Then, we calculate the PV of each simulation iteration using the WACC and take the average of these iterations to obtain an estimate of the expected PV. In the example project, the expected PV of the project without options is found to be \$48.39 million. Then, as the next step, cash flow payout rates are estimated for each time period using (2.1) and (2.2).

Period	δ_t	**Period**	δ_t	**Period**	δ_t	**Period**	δ_t
1	0.027	6	0.136	11	0.183	16	0.246
2	0.107	7	0.148	12	0.162	17	0.302
3	0.128	8	0.152	13	0.18	18	0.369
4	0.132	9	0.157	14	0.195	19	0.53
5	0.139	10	0.169	15	0.218	20	1

Table 2.6: Cash Flow Payout Rates

The PV and the cash flow payout rates are used to estimate the volatility of the GBM approximation. First, using this PV and the cash flow payout rates, and assigning an arbitrary value for volatility, a GBM approximation is modeled. Then, we search for the volatility that best mimics the uncertainty in the original SD model as described in Section 2.4. We use the risk analysis software @*Risk* in simulating the GBM process and estimating the cash flow percentiles. We find that when $\sigma = 0.06$, the GBM approximation mimics the cash flow distributions given by the original SD model quite closely (Figure 2.7). In this case, we limited the search to a predetermined set of candidate values for volatility; however, when more precision is required one can solve a stochastic optimization problem to determine the volatility that minimizes the difference between the original and the approximated cash flows.

As the next step (Step 7), we calculate the parameters of the binomial

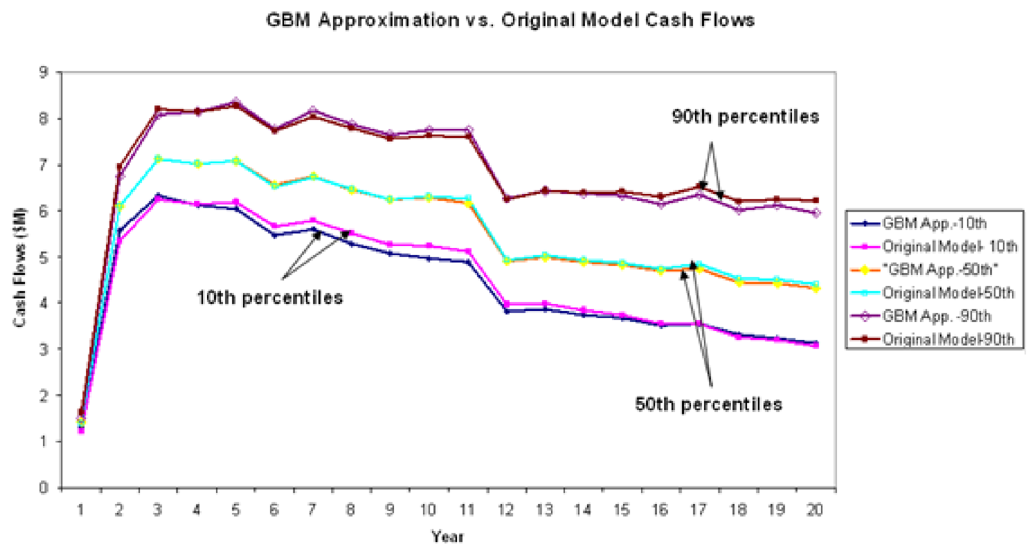


Figure 2.7: GBM Approximation vs. the Original Model Cash Flows

approximation of this GBM diffusion process. We set $\Delta t = 1$ year and assuming a 5% risk free rate, obtain $u = 1.06$, $d = 0.94$ and $p = 0.901$. Finally, the initial value V of the project is approximated by the PV of the project without options, \$48.39 million. Figure 2.8 depicts the first four periods of the binomial tree.

The last step is adding the options to the decision tree by using decision nodes (for details, see Brandao et al. 2005). For the example project, the cash flows change proportionally with the changes in the capacity. For example, the option to expand by 25 MW ("Expand Low") increases the capacity by 62.5%. The revenues from electricity generation are proportional to the capacity; hence, the increase in revenue when the option is exercised can be modeled by simply increasing the cash flows by 62.5%. Yet, this scheme does

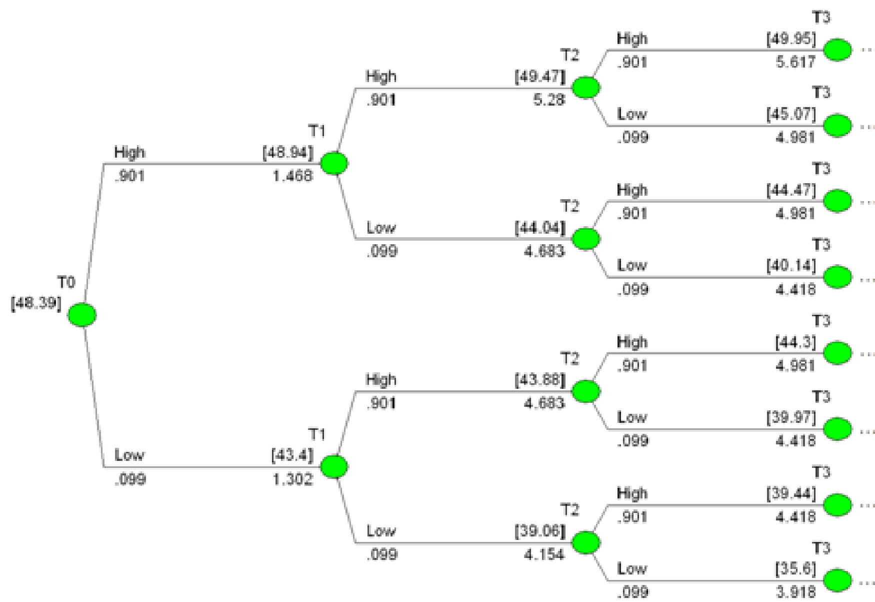


Figure 2.8: Binomial Tree Model of Value of Project without Options

not directly allow for incorporating the uncertainty in the cost of capacity (the learning curve uncertainty) because the cost of capacity is a one time payment at the exercise time of the option, which does not affect the cash flow stream afterwards. To handle this issue, we modeled the learning curve uncertainty as a *private risk*.

In many projects, there are project-specific risks that cannot be hedged by trading securities, such as technological risks. In view of that, real options studies make the distinction between public (market-priced) risks and private (project-specific) risks (Smith and Nau 1995). When the investment under concern is dominated by private risks, dynamic programming based approaches (such as decision tree analysis) should be preferred rather than the traditional option pricing techniques that were developed mainly for market-priced risks (Dixit and Pindyck 2001). Many projects have both kinds of risks, though; in which case the recommended strategy is to separate the public and private risks and use the appropriate risk adjustment for each one (Borison 2005, Brandao et al. 2005, Smith and Nau 1995).

Fortunately, treating different types of risks separately is straightforward once a decision tree is created. Public and private risks may be represented with separate chance nodes. Risk-neutral probabilities are used for the former and subjective probabilities are used for the latter. For the example project, the cost of capacity is discretized so that at any period t, it is *high*, *nominal* or *low*. The values for each of these branches were obtained by discretizing the Monte Carlo simulation data for the cost of capacity. This is done the same way as the cash flow distributions are discretized in the SD-based decision tree algorithm: A three-point bracket median method is used while preserving the path-dependence of the learning curve uncertainty by carefully computing the conditional probabilities of the branches. For example, chance nodes *HighC*, *NominalC* and *LowC* in Figure 2.9 discretize the learning curve uncertainty for the first period. Note that the uncertainty in the cost of capacity does not affect the volatility of the subsequent cash flows.

The decision tree was built and solved using DPL (Figure 2.9). The expected PV of the project is estimated to be \$57.48 million. The optimal policy suggests that the investment should be undertaken and it is optimal to expand high afterwards. Note that the SD-based decision tree approach results

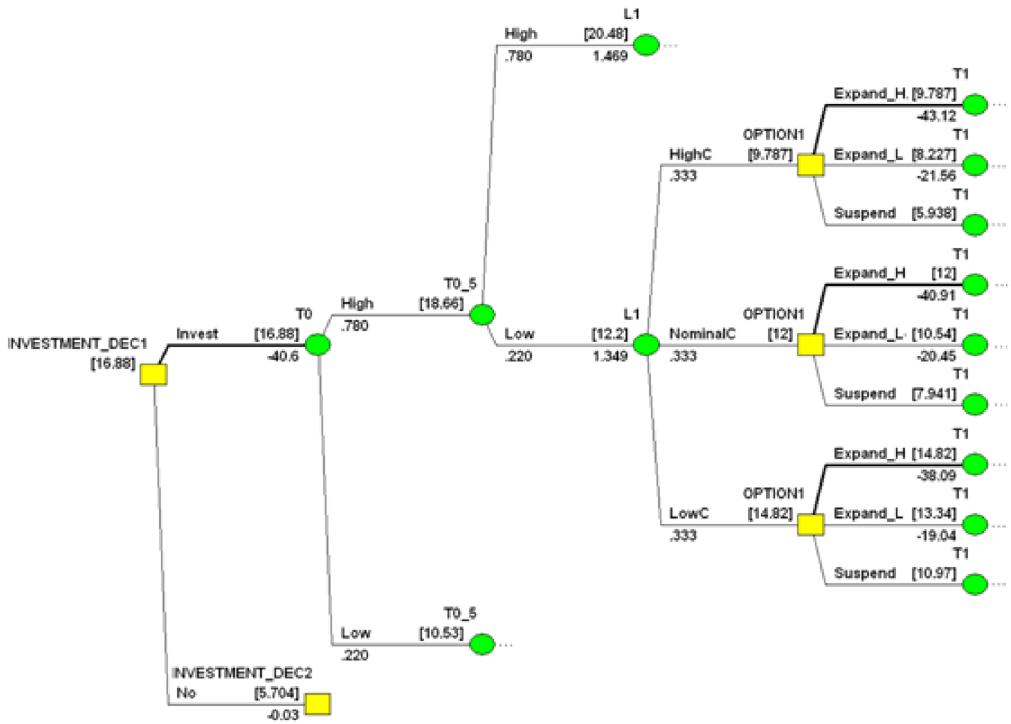


Figure 2.9: The Optimal Policy

in an expected PV of \$55.386 million and suggests investing immediately and expanding afterwards, the amount depending on the cash flow realization.

2.6 Discussion

Smith (1999) makes the observation that when evaluating risky projects, there has existed a fundamental trade-off between what he terms "detail complexity" and "dynamic complexity". He suggests that financial theory has tended to sacrifice detail complexity, the fidelity of a model at a detailed level, to better capture market information to appropriately discount risky cash flows. On the other hand, he suggests that decision analysis has often focused on detail complexity at the expense of keeping some model dynamics unrealistically simple, for example by using a single risk-adjusted discount rate to value future cash flows even if the risks of these cash flows change when different project options are selected. In this paper, we have attempted to show that by using a system dynamics model as an input to evaluating managerial flexibility it is possible to improve this trade-off between dynamic and detail complexity.

In particular, the diffusion approximation approach relies on building an SD model of the project to model the project uncertainty and a binomial tree approximation of the uncertainty to employ risk neutral valuation. The former brings high fidelity to model details due to the unique capabilities of SD in modeling complex feedback systems, whereas the latter avoids the need to estimate a risk-adjusted discount rate for the project with options ensuring better fidelity to dynamic complexity compared to traditional decision tree methods. Hence, the algorithm potentially improves the trade-off between dynamic and detail complexity especially for evaluating projects whose viability is determined by the interaction of stochastic processes within a complex nonlinear feedback structure.

For the example project, the SD-based decision tree approach and the diffusion approximation approach yielded similar results. However, this will not always be the case. The diffusion approximation algorithm overcomes the major flaw of the SD-based decision tree approach, which is the use of the same risk-adjusted discount rate for the project with and without options regardless of the changing risk character. In general, for projects with long lives or for projects whose risk-profiles change significantly, the errors caused by using the wrong discount rate are magnified, and the differences between the two approaches become larger (Smith and McCardle 1998, Teisberg 1995).

Nonetheless, diffusion approximation approach has some limitations of its own. The algorithm makes two strong assumptions. One is that the cash flow of the project without options can be represented by a GBM (or other similarly tractable) stochastic process. This implies that the algorithm will only yield a reasonably accurate valuation if each period's cash flow from the project without options has an approximately lognormal distribution. Fortunately, because of the numerous influences upon the individual cash flows that one can represent in a system dynamics model, this is often the case.

A second restriction is that the diffusion approximation approach assumes the cash flows of the project with an option are proportionate to the cash flows of the project without the options. For example, if the firm exercises an expansion option to increase its capacity by $x\%$, the revenues should change by a linear function of x. If this is not the case, then the diffusion approximation approach may not be an appropriate modeling approach. In such cases, the SD-based decision tree approach might be a better alternative since it provides a greater fidelity to the details of the modeled project and is not subject to the assumptions associated with the use of the GBM diffusion approximation. Yet, for a reliable valuation the analyst needs to make sure that the valuation results are not highly sensitive to the choice of discount rate. If project options do not create a significant change in the risk-profiles of their associated cash flows, the SD-based decision tree approach should provide an accurate estimate of the project value.

2.7 Summary and Conclusion

The viability of many projects is determined by a number of stochastic dynamic processes interacting within a complex, nonlinear feedback structure. For example, the value of alternative energy technology projects is largely determined by fossil fuel prices and capital equipment costs, both of which are highly stochastic and tied to complex feedback structures. System dynamics is a methodology developed to analyze and manage complex feedback systems. Consequently, it is powerful in handling nonlinearity and path-dependence unlike the traditional methods used in representing stochastic dynamic processes.

The major contribution of this paper is in demonstrating how to evaluate projects with uncertainties embedded in feedback loops by using a real options valuation approach and stochastic system dynamics models. We propose two methods, which we termed the *SD-based decision tree approach* and the *diffusion approximation approach*, the former being based on traditional decision tree analysis and the latter being based on modern finance theory. Both methods transform data obtained from Monte Carlo simulations of an SD model into an approximate decision tree, taking advantage of the complementary strengths of system dynamics and decision trees.

Chapter 3

Platform Performance Investment in the Presence of Network Externalities

3.1 Introduction

Video-game industry is one of the fastest growing industries in the entertainment sector. It has already overtaken Hollywood box office and is expected to overtake music industry in North America and UK (EMA 2008, Foster and Ahmad 2008, Cheng 2007, Money 2007). This almost recession-proof growth trend is partly made possible by the 7th generation game consoles' success in unlocking the market's potential. The success story of Nintendo Wii is particularly interesting and has provided the motivation behind this study. Nintendo Wii is released later than the other 7th generation game consoles, Sony's PS3 and Microsoft's Xbox, which may imply a significant disadvantage in markets that exhibit network effects (Arthur 1989, Katz and Shapiro 1994). Furthermore, the processing and graphical capabilities of Wii fall short of the bar set by PS3 and Xbox (Allen 2006). Nonetheless, Wii has the biggest installed base among the three (VG Chartz Home Page 2009).

Wii's story stirred a lot of discussion advocating different reasons to account for this success. One popular explanation has been Wii's lower price,

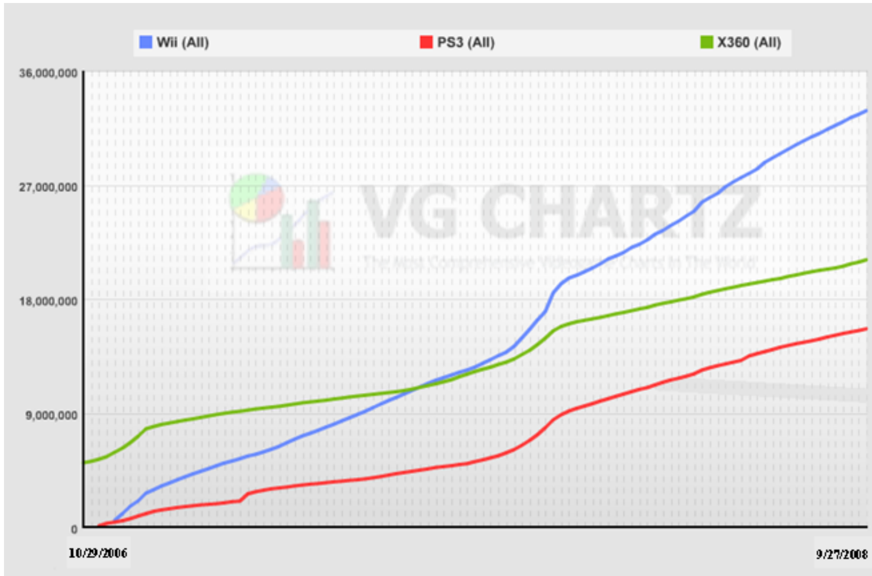


Figure 3.1: 7th generation video-game consoles cumulative sales. Taken from vgchartz.com

which definitely gives it an edge against PS3. Yet, Wii's price advantage over Xbox is not as significant. Given the early entry advantage of Xbox, pricing alone is unlikely to explain Wii's success. Another success factor is Wii's unique motion-sensitive remote control, Wii Remote. Wii Remote creates a unique gaming experience that is more welcoming to novice gamers because of its intuitiveness. Yet, most of the games developed for Wii still use the traditional joystick controls instead of the unique remote, thus the remote alone fails to fully account for Wii's story. It has also been speculated that Nintendo starved the market deliberately before the holiday seasons which created a buzz and increased the desirability of owning a Wii (McIntyre 2008, Kane and Wingfield 2007). Similar claims have also been made for earlier generation video-game consoles and the potential benefits of a deliberate shortage have been well discussed (Brandenburger and Nalebuff 1996). However, even if Wii benefited from a deliberate shortage, Xbox 360 had a similar "strategy" as well (Wingfield and Guth 2005). Hence, a shortage induced fad does not quite explain how the Wii could overtake Xbox 360.

One particular reason behind Wii's success provided the motivation for this paper. Due to its lower processing and graphical capabilities that are not much different than 6th generation technology, it is significantly cheaper to develop games for Wii (Leheng 2006, Sinclair 2006), making the platform attractive for game developers. Indeed, in spite of its late release, the Wii has more game titles compared to its competitors (VG Chartz Game Database 2009). Many gamers seek a variety of games to play; hence the availability of a variety of game titles makes a game console more attractive for the gamers.[1] Likewise, a bigger installed base attracts game developers in the anticipation of higher sales. These cross-network effects largely determine the fate of a game console and they have worked out quite favorably for the Wii [2].

Wii's story revealed that even though a game console with a higher performance such as one with better graphical capabilities is attractive to

[1]Importance of software is well recognized in the industry. In an interview, Reggie Fils-Aime, Nintendo's current U.S. president states that Nintendo is putting a lot of effort in encouraging publishers to make games for the Wii to ensure a steady flow of diverse games saying " Price cuts are a short-term incentive. In the long run, you need software to excite people. In order to achieve high levels of sales of hardware, we need all genres in the market" (Kane and Wakabayashi 2009).

[2]Brandenburger and Nalebuff (1996) provide an excellent account of how Nintendo created a similar "virtuous circle" with its Nintendo Entertainment System in late 1980s.

gamers, the technology behind high performance may make it more costly to develop games and hinder game development. Furthermore, gamers' interest in further improved graphics might be much less than what is assumed by the console manufacturers (Sheffield 2008). In other words, console manufacturers might be overinvesting in the core performance of their platform. Indeed, Van der Rhee et al. (2009) discuss that steepening the performance treadmill is not always the best strategy in a competitive market. But how should a console manufacturer balance the gamers' preferences with game developers' needs to fully benefit from cross-network externalities?

A similar question can be posed for other *two-sided markets*. Two-sided markets are markets in which one or several platforms facilitate interactions between two distinct group of users whose interactions are subject to cross-network effects (Eisenmann, Parker and Alstyne 2006, Rochet and Tirole 2006). Examples can be found in various industries such as cell phones, credit cards, online dating services, travel reservation services, and shopping malls. In many cases platform sponsors face a trade-off between developing a high performance platform that matches the end-user's preferences and sacrificing some of those preferences in exchange for improved or less costly 3rd party development capabilities.

In this paper, we study the performance[3] choice of two competing plat-

[3]We use the term "performance" to represent a vertically differentiated dimension of quality (Mussa and Rosen 1978). For video-game consoles, better graphics and better processing capabilities imply higher performance.

form sponsors where higher performance makes third party content development more costly, either directly as in the case of video game platforms or indirectly by taking away resources that could have been used on improving third party development capabilities. We consider two different games based on the order of market entry: a simultaneous-move game where platforms enter the market at the same time and a sequential-move game where there is a leader and a follower in the market. We provide insights on the optimum investment in platform performance and show that conventional wisdom of product development may be misleading in the presence of strong cross-network externalities.

In particular, we characterize when a platform market is *content-driven* versus when it is *performance-driven* and discuss how conventional wisdom could mislead a platform sponsor especially in content-driven markets. For instance, in a one-sided market if competition between two platforms intensifies, one would expect that at equilibrium there will be more investment to increase the performance of the product. However, we show that platform sponsors may be better off decreasing the investment in platform performance to provide greater content availability instead. Likewise, an increase in end-users' interest in platform performance may still trigger a less aggressive investment in platform performance if the market is content-driven. More importantly, we show that when platforms are price-takers, a platform with a lower performance can capture a bigger market share on both sides of the market, similar to the success of the Wii. In other words, making the highest investment in technology is not the only way to be the winning platform. In certain markets, the key to success is in better mobilization of third party developers.

We also observe that when platforms are able to determine the end-user price in addition to the platform performance, two-sidedness presents itself in the interplay between these decisions. For example, when certain market parameters change, it may be optimal to reduce the investment in the performance of the platform and yet increase the price at the same time. With a slightly lower performance, a platform may become more attractive to content developers which can indeed justify the higher end-user price in a content-driven market.

Throughout the paper, we use the video-game industry as an illustrative example; yet the model is applicable to other two-sided markets with similar features. For example, providers of smartphones face a similar trade-off of balancing the needs of application providers and end-users.

The remainder of the paper is as follows. Section 3.2 reviews the related literature. In Section 3.3, we develop a mathematical model of two competing platforms that operate in a two-sided market. In Section 3.4, we consider platforms that enter the market simultaneously. First part of this section assumes that the platform sponsors are price-takers (PT) whereas in the second part, this assumption is partially relaxed by allowing the platform sponsors to set the end-user price in addition to the platform performance (PS). In Section 3.5, we briefly present the sequential-move counterpart of the same analysis. Then, we discuss the limitations of our model and possible extensions in Section 3.6. Finally, Section 3.7 summarizes the results and concludes the paper.

3.2 Literature Review

There is a burgeoning stream of research pioneered by Parker and Van Alstyne (2000, 2005) and Rochet and Tirole (2003) on the economic theory of two-sided markets which focuses on the unique features that set these markets apart from traditional products and services. In particular, two-sided markets exhibit a special form of indirect network effects (Katz and Shapiro 1985, Liebowitz and Margolis 1994) where the number of users on one side of the market depends on the number of users on the other side. For example, video game developers will develop games only for platforms that have a sufficiently big installed base of gamers. Likewise, all else being equal, gamers prefer platforms that provide a greater variety of games. Due to the presence of these *cross-network effects*, a platform sponsor may subsidize one side of the market in order to attract the other side (Eisenmann et al. 2006). The growing literature on two-sided markets has mostly focused on these novel two-sided pricing strategies adopted by platform sponsors (Parker and Alstyne 2000, Parker and Alstyne 2005, Armstrong 2006, Caillaud and Jullien 2003, Hagiu 2006, Rochet and Tirole 2003, Rochet and Tirole 2006).

Specifically, Parker and Van Alstyne (2000, 2005) characterize the pricing structure of a monopolist platform and show that either side of the market may be subsidized depending on the relative network externality benefits. Rochet and Tirole (2003) develop the two-sided pricing strategies for a wide range of governance structures including competing profit and non-profit platforms. Similarly, Armstrong (2006) analyze both monopolistic and competing platforms, and show that the pricing structure depends on the relative strengths of cross-side network effects, the fee structure and whether the agents are able to join more than one platform.

Two-sided markets literature has provided valuable insights on the role of pricing strategy in capturing demand on both sides of the market. Yet, very little has been done to explore the use of non-price controls. There is a recent body of work discussing the non-price levers platform sponsors can use to create more attractive bundles for the end-users (Gawer and Cusumano 2002, Boudreau and Hagiu 2008, Parker and Alstyne 2008, Eisenmann, Parker and Alstyne 2007). However, the role of platform characteristics such as platform features and performance has been largely ignored. On the other hand, there is a substantial body of research on product development that investigates these issues for traditional products and services. For instance, marketing based studies in the product development literature have extensively analyzed how to determine the target values of attributes of a product (see Green and Srinivasan 1990 for a review) whereas operations management based studies have focused on determining the design parameters that will optimize product performance (see Papalambros 1995 for a review). For a broader view of the product development research, we refer the reader to the survey papers in this area (Krishnan and Ulrich 2001, Shane and Ulrich 2004). Our paper brings together product development and two-sided markets literatures by analyzing a platform sponsor's investment in platform performance as a non-price control to "get the two sides on board".

There are very few studies in the two-sided markets literature that focus on platform design issues such as platform quality and features. Bhargava and Choudhary (2004) analyze the product line design (or "versioning") problem of an information intermediary showing that when the buyers have constant marginal valuations for the service quality, versioning is optimal. A study that is closer to this paper is by Zhu and Iansiti (2009), who consider two platforms, an incumbent and an entrant, competing on the basis of platform quality and installed base. The authors analyze a dynamic game and show that installed base does not necessarily present barriers to entry. Our paper, on the other hand, mainly focuses on how investment in platform performance differs from product development strategies in the absence of cross-network effects. Further, in Zhu and Iansiti (2009) developers' cost is independent of the platform quality, whereas the (direct or indirect) adverse effect of high platform performance on development costs and the resulting performance trade-off is at the core of our model.

3.3 The Model

This section describes the two-sided market model that we use to analyze platform performance investment decisions in the presence of cross-network externalities. Throughout the paper, we draw on the video-game industry as an illustrative example. In line with the previous literature on two-sided markets, we use the terms end-user, content developer and platform sponsor to correspond to gamers, game developers and game console providers respectively.

We assume that content developers may choose to affiliate with more than one platform, i.e. "multihome", whereas end-users purchase a single game console; i.e. "singlehome". The assumption that content developers may multihome is based on the current trend of publishing multi-platform games, which is anticipated to prevail due to the growing cost of game development (Reisinger 2009, Nutt 2008). End-users, on the other hand, are assumed to singlehome since it is not very common to own more than one platform of the same generation. Further, it can be argued that as more developers produce multi-platform content, fewer end-users will choose to multihome; because the amount of content that is exclusive to a platform outside the end-user's access will be limited (Dutka 2009). In general, it is not common to see platforms where both sides multihome, since if every agent on one side of the market joins all platforms there is no incentive for any agent on the other side of the market to join more than one platform (Armstrong 2006).

Because they are able to multihome, content developers decide whether or not to join a particular platform independently from their participation decision for the other platform (ignoring any budget constraints). End-users, on the other hand, need to decide which platform to join, creating competition between the platforms to attract them. End-users that choose platform i pay p_i

Table 3.1: Notation

	Description
ϕ_i	Performance of platform i
p_i	End-user price of platform i
D_i	End-user market share of platform i
N_i	Content developer market share of platform i
α_g	Content price
γ	Royalty per content sold ($\gamma < \alpha_g$)
β_g	Content development cost per unit performance
M	Fixed cost a content developer incurs for joining a platform
K	Platform development cost per unit performance squared
c	Marginal cost of platform production
α_c	End-users' utility from an additional unit of content available
β_c	End-users' utility from an additional unit of platform performance
t	Degree of platform differentiation on the end-user market

$(i = 1, 2)$ to purchase the platform whereas content developers pay a royalty of γ per content sold. We assume that there is perfect competition among content developers, hence contents cost α_g regardless of the developer. Indeed, in the video game industry, game prices are more or less the same at the release date irrespective of the developer and the console they are developed for. For the ease of exposition, it is also assumed that γ and α_g are the same for both platforms. Hence, the explicitly modeled dimensions of platform differentiation are platform performance ϕ_i and end-user price p_i $(i = 1, 2)$. Depending on these attributes, platform i gets D_i percent of the end-user market and N_i percent of the developer market $(i = 1, 2)$. The prospect of these market shares plays a major role in the platform choice of end-users and content developers.

The value an end-user obtains from purchasing a platform has two major components: platform performance and the availability of content for the platform. Each component has a certain weight in an end-user's decision. Let α_c be an end-user's utility from an additional unit of content available[4] and β_c be an end-user's utility from an additional unit of platform performance. The utility of end-users from joining platform i ($i = 1, 2$) is

$$u_i = \alpha_c N_i + \beta_c \phi_i - p_i \tag{3.1}$$

To determine the end-user market shares, we use the Hotelling's linear city model: Individual end-users are assumed to be uniformly located along a unit interval and the platforms are located at the opposite ends of the interval. Let t be the "transportation cost" parameter in the Hotelling model, which represents the degree of horizontal product differentiation between the platforms. Note that low t implies less product differentiation; thus a higher degree of competition. Without loss of generality, assume that Platform 1 is located at point 0 whereas Platform 2 is located at point 1. Accordingly, the net utility from joining Platform 1 for the end-user x located at $x \in [0, 1]$ is:

$$u_{xi} = \alpha_c N_i + \beta_c \phi_i - p_i - tx$$

By locating the marginal end-user that is indifferent between the two platforms and using the fact that end-users are uniformly distributed on a unit interval, the fraction of end-users that join platform i ($i = 1, 2$) can be calculated as:

$$D_i(\phi_i, \phi_{-i}, N_i, N_{-i}) = 1/2 + \frac{u_i - u_{-i}}{2t} \tag{3.2}$$

[4] We assume that α_c accounts for the cost of the content as well; i.e. $\alpha_c N_i$ is the net benefit from content availability

We normalize the market size to 1; hence D_i is equivalent to the number of end users joining the platform i.

We assume that a content developer's revenue has two components: For all developers, joining platform i generates a base revenue determined by the end-user market share of the platform, D_i, and the platform performance, ϕ_i, $(i = 1, 2)$. Further, each developer has an additional random component of revenue ϵ_j to account for the "hit factor": Content that meets user needs more successfully generates more revenues, such as "hit" games. From each unit of content sold, a content developer earns α_g but has to pay γ as royalty to the platform sponsor. We assume $\alpha_g \geq \gamma$ to guarantee that content developers have nonnegative margins. Under the assumption that end-users purchase every content developed for the platform[5], the revenue of content developer j for developing content for platform i, $R_{j,i}$,is given as follows:

$$R_{j,i} = D_i(\alpha_g - \gamma) + \epsilon_j$$

Content developers incur a development cost, which increases with the platform's performance. For simplicity, this cost is modeled as $\beta_g \phi_i$ where β_g is the cost of content development per unit performance. There is also a fixed cost M associated with developing content for a platform, independent from the platform's performance. For example, in the video-game industry, game developers need to purchase the software development kit before developing for

[5]This assumption can easily be relaxed without affecting the results qualitatively by assuming each end-user on average buys a certain fraction of the available content.

a specific platform. Further, hiring new employees or training the employees to develop for a new platform also results in a fixed cost. Content developers are assumed to be profit maximizers. Accordingly, by normalizing the opportunity cost to zero, it can be seen that a potential content developer for platform i ($i = 1, 2$) enters the market if $D_i(\alpha_g - \gamma) + \epsilon_j - \beta_g\phi_i - M \geq 0$. For simplicity, we assume that the random component in a content developer's revenue is uniformly distributed, $\epsilon_j \in U[0, 1]$. Hence, the maximum excess revenue a developer can get from the hit factor is normalized to 1. Further, we assume $M < 1$ on the grounds that the fixed cost of joining a platform is unlikely to exceed the most successful content's excess revenue. Under these assumptions, the marginal developer m, who is indifferent between developing and not developing content for platform i ($i = 1, 2$) is characterized with ϵ_m such that:

$$\epsilon_m = \beta_g\phi_i + M - D_i(\alpha_g - \gamma)$$

Accordingly, content developers with $\epsilon_j \geq \epsilon_m$ join platform i ($i = 1, 2$). We normalize the size of the content developer market to 1; hence the number of content developers who join platform i ($i = 1, 2$) as a function of performance decisions and the end-user market share is given by

$$N_i(\phi_i, D_i) = 1 - \beta_g\phi_i - M + D_i(\alpha_g - \gamma) \tag{3.3}$$

To calculate the end-user market share, we substitute (3.3) and (3.1) into (3.2) to get

$$D_i(\phi_i, \phi_{-i}) = 1/2\ (1 + \frac{(\phi_i - \phi_{-i})(\beta_c - \alpha_c\beta_g)}{t - \alpha_c(\alpha_g - \gamma)}) \tag{3.4}$$

Further substituting (3.4) into (3.3), we get

$$N_i(\phi_i, \phi_{-i}) = 1 - M - \beta_g \phi_i + \frac{\alpha_g - \gamma}{2}(1 + \frac{(\phi_i - \phi_{-i})(\beta_c - \alpha_c \beta_g)}{t - \alpha_c(\alpha_g - \gamma)}) \qquad (3.5)$$

The platform sponsors are profit maximizers. They have two sources of revenues: Revenue collected from the end-users who purchase the platform and the royalty revenue collected from content developers from each content they sell to the end-users. There is a fixed cost of developing the platform, which is assumed to be a convex increasing function of platform performance given as $K\phi_i^2$. For ease of exposition, the marginal cost of production, c, is assumed to be constant[6] and identical for both platforms. Accordingly, platform sponsor i's $(i = 1, 2)$ profit function, π_i is as follows:

$$\pi_i(\phi_i, \phi_{-i}) = (p_i - c)D_i(\phi_i, \phi_{-i}) + \gamma D_i(\phi_i, \phi_{-i})N_i(\phi_i, \phi_{-i}) - K\phi_i^2 \qquad (3.6)$$

where $D_i(\phi_i, \phi_{-i})$ and $N_i(\phi_i, \phi_{-i})$ are given by (3.4) and (3.5) respectively.

The analysis of the market equilibrium that arises from this model depends on the timing of the decisions and the pricing power of the platform sponsors. In terms of the timing of the decisions, platforms may make their decisions simultaneously (a simultaneous-move game) or there may be a leader and a follower in the market (a sequential-move game). In terms of the pricing power, the platforms may be price-takers or price-setters. We start with analyzing the simultaneous move game for price-taker and price-setting platforms,

[6] All of our results still hold qualitatively if the marginal cost is linearly increasing in platform performance; i.e. $c\phi_i$ where c is a constant.

both of which yield analytically tractable results. In Section 3.5, we briefly look at the impact of a leader-follower structure for which we mostly rely on numeric analysis.

3.4 Analysis: Simultaneous-move game

In this section, we consider two competing platforms that enter the market simultaneously; in other words both platforms make their decisions without observing the competitor's decisions. We analyze two main scenarios: A duopoly of price-taker platforms versus a duopoly of platforms that can set the end-user price in addition to platform performance.

3.4.1 A duopoly of price-takers

Analyzing price-taker platforms enables us to focus on platform performance as a tool to capture demand in a two-sided market. By singling out the performance choice, the interplay between end-users' preferences and content developers' constraints becomes more evident. The assumption of price-taker platforms fits the video-game industry reasonably well. Largely because of the price sensitivity of gamers, console providers almost behave as price-takers. Harvard Business School case for Sega Enterprises provides an interesting account on the degree of gamers' price sensitivity: For Sega, a price choice that exceeds the post-price war bottom by 25% is estimated to drop the inclination to buy among the target market by 60-70%. (Thomke 1999). Hence, gamers' price sensitivity puts a strict upper limit on the console price. On the other hand, the significant costs involved in the development and manufacturing of game consoles put a strict lower limit on the price, making the price-taker assumption a reasonable fit.

Consider two competing price-taker (PT) platforms. Specifically, platform sponsors commit to the end-user price p_i in advance, which leaves platform performance as the only lever to capture demand on both sides of the market. Accordingly, a platform sponsor's decision problem is as follows:

$$\begin{aligned} &\max_{\phi_i \geq 0} (p_i - c) D_i(\phi_i, \phi_{-i}) + \gamma D_i(\phi_i, \phi_{-i}) N_i(\phi_i, \phi_{-i}) - K\phi_i^2 \\ &s.t. \;\; \phi_i \geq 0 \end{aligned} \tag{3.7}$$

We first analyze a benchmark case where platforms commit to the same end-user price. This special case is tractable, which enables us to provide analytical results. Then, we provide some insights for the general case where this assumption is relaxed.

3.4.1.1 A Duopoly of Price-takers with Symmetric Platforms

In this section, we analyze two competing platforms that enter the market simultaneously and commit to the same end-user platform price; i.e. $p_1 = p_2 = p$. In other words, we consider a duopolistic platform market where the two firms commit to equal prices and compete on performance. Accordingly, we solve (3.7) for both platforms simultaneously and obtain a symmetric equilibrium where both platform sponsors choose the following performance level and split the market equally.[7]

$$\phi_1^{PT} = \phi_2^{PT} = \phi^{PT} = \frac{m_p v_\phi - \gamma(t\beta_g - \beta_c(\alpha_g - \gamma))}{\beta_g \gamma v_\phi - 4K\chi} \tag{3.8}$$

where $v_\phi = \beta_c - \alpha_c\beta_g$, $\chi = \alpha_c(\alpha_g - \gamma) - t$ and $m_p = (p - c + \gamma(1 - M))$.

v_ϕ provides a measure of the market value of platform performance. When v_ϕ is negative end-users tend to value content availability more than they value platform performance whereas content developers suffer from a high content development cost per unit performance. In that case, a high platform performance is not appreciated in the market; hence the value of platform performance is low. On the other hand, when v_ϕ is positive, content development cost per unit performance is not prohibitive whereas end-users' utility from platform performance is significant; which implies a high market value of platform performance. The competition factor χ, on the other hand, is a rough measure of the strength of competition relative to the strength of cross-network effects. When χ is negative, there is not much competition between the two platforms in the end-user market and the networks effects are weak whereas a positive χ implies a highly competitive market with significant network effects. Finally, m_p provides an aggregate measure for the platform sponsors' margin. We assume that $m_p \geq 0$ for the platform sponsors to stay in business. Even if the end-user price is below marginal cost, the royalty revenue can compensate to result in a positive margin. This allows our framework to

[7]The expression in (3.8) is the unconstrained performance decision; i.e. the constraint $\phi \geq 0$ is not enforced. To ensure $\phi^{PT} \geq 0$, we restrict the parameter space, as explained in Appendix B.1

account for the widely seen phenomenon of pricing below marginal cost in one side of the market (Parker and Alstyne 2005)

From Lemma B.1.1 (in Appendix B), market participation in the developer side occurs only when v_ϕ and χ have opposite signs. In other words, under our assumptions there are two types of markets in which the platforms can profitably do business. In the first type, the *content-driven market*, market value of platform performance is low, yet the competition factor is high. This implies that end-users highly value content availability and that the performance difference between the two platforms is less consequential for their decision. In the second type, the *performance driven market*, end-users' focus shifts to the performance of the platform. Additionally, the two platforms are differentiated enough to appeal to different segments of end-users alleviating the intensity of competition. For the video-game industry, the first type would correspond to a market dominated by "casual gamers", whereas the second would be a predominantly "hard-core gamers" market. Even though both terms are loosely defined, it is generally assumed that hard-core gamers appreciate the graphical and processing capabilities of a game console to a better extent compared to casual gamers. On the other hand, casual gamers are typically interested in casual games that are quick to access, easy to learn, and that do not require gaming expertise or regular time commitment to play (Casual Games Association 2007).

The performance investment at equilibrium presents some interesting features that would not be observed in the absence of cross-network effects.

We uncover some of these features in Proposition 3.4.1 through comparative statics analysis.[8]

Proposition 3.4.1. *At equilibrium, the platform performance ϕ^{PT} is:*

i) decreasing in the end-users' utility from platform performance, β_c, if market value of performance v_ϕ is negative and competition factor χ is positive.[9]

ii) decreasing in the degree of competition among the platforms in the end-user market, if market value of performance v_ϕ is negative and competition factor χ is positive.

iii) decreasing in the content developer's revenue per content, α_g, if market value of performance v_ϕ is negative and competition factor χ is positive.

iv) increasing in the end-users' utility form content availability, α_c, if $t\beta_g - \beta_c(\alpha_g - \gamma) \leq 0$.

In a one-sided market if consumers highly value the performance of a product, competition would drive firms to offer higher performance. Proposition 3.4.1 i) shows that this is not necessarily the case in two sided markets. In particular, platform sponsors may choose to reduce their investment in

[8] All proofs are provided in the Appendix B.

[9] Throughout the paper unless otherwise specified, we mean increasing and decreasing in the weak sense.

platform performance in response to increasing end-user utility from performance. The fundamental reason behind this counterintuitive result is the fact that platform performance and content availability act as substitutes for the end-users. Consider a highly competitive platform market characterized by end-users with high preference for content availability and content developers that have a high development cost per unit performance; in other words, consider a content-driven market. In such a market, a platform sponsor might be better off by decreasing the investment in the performance of the platform when end-users' utility from performance increases. Due to the high content development cost per unit performance, a slight decrease in the platform's performance may have a big impact on attracting new content developers, which ultimately attracts end-users as they enjoy a high utility from content availability.

Part ii) presents a similar result for the relation between the degree of competition and the platform performance. In the absence of cross-network effects, if competition between two firms intensifies, one would expect a higher investment in performance at equilibrium. However, in a two-sided market platform performance decreases as the competitiveness in the end-user market intensifies if the competition factor is high but the market value of platform performance is low. Similar to part i), this stems from the substitution effect between content availability and platform performance. When the market value for platform performance is low, instead of providing a platform with higher performance as a response to increasing competition, the platform sponsors may be better-off investing slightly less in performance and getting content developers on board.

Part iii) can be interpreted in a similar fashion. For a two-sided market, the investment in platform performance may drop in response to an increase in the content price if the market is content-driven. Even though a higher content price could easily demand an increase in platform performance, when end-users value content availability more than they value the performance of the platform, a lower performance choice that lowers the content development cost and attracts more content developers is more desirable for the platforms.

In the first three parts of Proposition 3.4.1, we observe that conventional wisdom would mislead a platform sponsor in a content driven market. In part iv), conventional wisdom fails if the end-users have a high valuation for platform performance, as in the case of hard-core gamers. When the end-users' utility from content availability increases, a platform sponsor may choose to invest less in the performance of the platform for the sake of attracting more content developers, which attracts the end-users. However, this may not be the the best response in a highly competitive platform market characterized by end-users with a high utility from platform performance and content developers that enjoy a high margin in addition to a development cost structure that is not significantly affected by platform performance. In such markets, higher performance makes the platform significantly more attractive for the end-users whereas it does not increase the development cost of the content developers as much; hence the platform sponsor is better off increasing the platform performance when end-users' utility from content availability increases.

Overall, Proposition 3.4.1 implies that a firm may overinvest or under-invest in platform performance should the cross-network effects be ignored or miscalculated. Note that except for part iv) this happens in a content-driven market.

3.4.1.2 A Duopoly of Price-takers with Asymmetric Platforms

In this section, we relax the assumption that the two platforms commit to the same end-user price. This general case is not tractable, even though closed-form solutions for (ϕ_1, ϕ_2) still exist. Yet, we are able to provide some insights through numeric analysis.

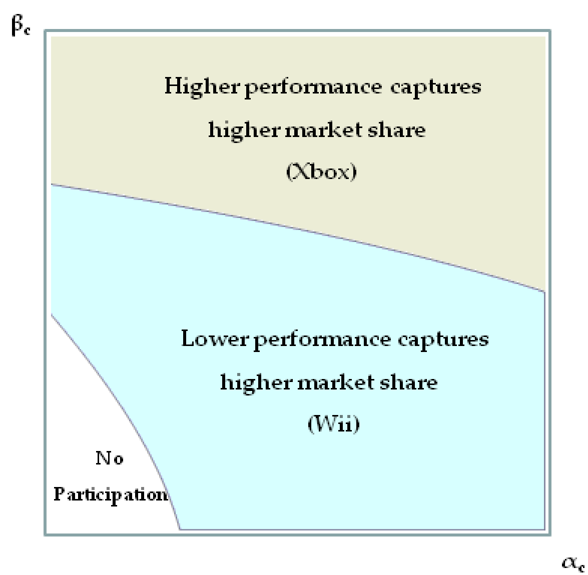


Figure 3.2: An example market segmentation under high competition

An asymmetric duopoly presents a richer variety of market segmentation scenarios among which exist ones that mimic Wii's success story. In other words, we observe that the platform with a lower investment in performance may be the market leader. The following example represents a highly competitive end-user market where Platform 1 commits to a slightly lower end-user price compared to Platform 2. The bottom left corner of Figure 3.2 shows that when end-user utility from content availability (α_c) and end-user utility from platform performance (β_c) are both low, neither platform succeeds in attracting users. In the "Wii" region, Platform 1 has a lower performance, yet still captures a bigger market share on both sides of the market. Note that this is most likely to happen in a market characterized by high end-user utility from content availability and a relatively low end-user utility from platform performance. The "Xbox" region, on the other hand, shows where the more expensive and higher performance platform becomes the market leader. Note that this requires a sufficiently high utility from platform performance. This market segmentation scenario shows us that making the biggest investment in the platform technology does not necessarily bring market leadership. In content driven markets, the key to capturing demand is triggering third party development.

To complete the analysis, we look into the relation between the committed price p_i and the equilibrium performance. We first observe that in line with conventional wisdom if a platform sponsor commits to a higher end-user price, it offers a higher performance platform. The effect of competitor's end-user price on a platform's performance choice turns out to be more subtle though.

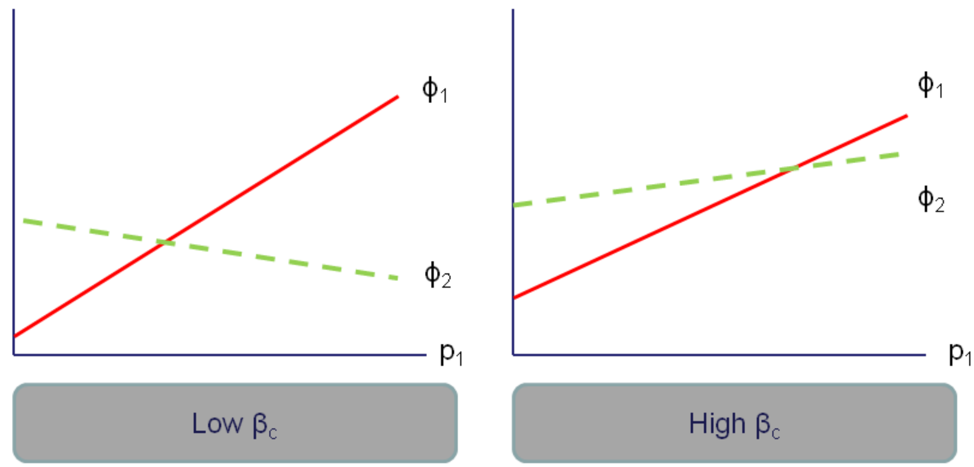


Figure 3.3: Relationship between platform performance ϕ_i^{PT} and the competitor's end user price p_{-i} depends on the market value of performance, v_ϕ

That is best explained with an example. In Figure 3.3, the horizontal axis is the price of Platform 1 and the vertical axis represents the platform performance. On the left panel, it can be seen that the performance of Platform 2 decreases as the price of Platform 1 increases, while on the right panel it increases as the price of Platform 1 increases. The settings on the two panels are identical except for the end-user utility from the platform performance and the end-user utility from content availability. When the end-users' utility from platform performance is higher compared to their utility from content availability, for example in a performance-driven market, the platform sponsor is better off increasing the performance as a response to increasing end-user price of the competitor. On the other hand when the end-users' utility from platform performance is low, for example in a content-driven market, the platform sponsor is able to have a less aggressive investment in platform performance taking advantage of the competitor's increased price.

3.4.2 A Duopoly of Price-setting Platforms

In this section, we partially relax the price-taking assumption by allowing the platform sponsors to determine the end-user price. However, royalty collected from the content developers is still exogenous to the model, since it is not possible to obtain closed form solutions when platform sponsors determine the two-sided pricing strategy and platform performance simultaneously. Accordingly, a price-setting platform sponsor's decision problem is as follows:

$$\begin{aligned} &\max_{\phi_i \geq 0; p_i} (p_i - c) D_i(\phi_i, \phi_{-i}) + \gamma D_i(\phi_i, \phi_{-i}) N_i(\phi_i, \phi_{-i}) - K\phi_i^2 \\ &s.t.\ \ \phi_i \geq 0 \end{aligned} \tag{3.9}$$

When two symmetric platform sponsors determine the end-user price and the platform performance simultaneously, the price-setting (PS) equilibrium is symmetric with both platforms setting the end-user price and platform performance specified in Lemma 3.4.2

Lemma 3.4.2. *If $v_{PS} = \beta_c - \beta_g(\alpha_c + \gamma) \geq 0$, platforms play the high perfor-*

mance equilibrium given by

$$
\begin{aligned}
\phi_H^{PS} &= \frac{\beta_c - \beta_g(\alpha_c + \gamma)}{4K} \\
p_H^{PS} &= \beta_g \gamma \frac{\beta_c - \beta_g(\alpha_c + \gamma)}{4K} + t + c - \gamma(1 - M + \alpha_g - \gamma) - \alpha_c(\alpha_g - \gamma) \\
&= \beta_g \gamma \phi_H^{PS} + t + c - \gamma(1 - M + \alpha_g - \gamma) - \alpha_c(\alpha_g - \gamma)
\end{aligned}
$$

Otherwise, the low performance equilibrium (L) is played with zero investment in platform performance:

$$
\begin{aligned}
\phi_L^{PS} &= 0 \\
p_L^{PS} &= t + c - \gamma(1 - M + \alpha_g - \gamma) - \alpha_c(\alpha_g - \gamma)
\end{aligned}
$$

Hence, when two competing symmetric platforms set the end-user price and platform performance simultaneously, end-users' utility from platform performance must be relatively high and content development cost per unit performance must be relatively low for the platforms to positively invest in platform performance. On the other hand, if end-users tend to value content availability more than they value platform performance and content developers suffer from a high content development cost per unit performance as well as a high royalty rate, a high platform performance is not appreciated in the market which induces the platforms to play the low performance equilibrium.

Looking at closely to these equilibrium decisions show that the equilibrium performance level behaves quite differently compared to the price-taker equilibrium, as summarized in Proposition 3.4.3.

Proposition 3.4.3. *In a price-setting duopoly where platform sponsors set the end-user price and platform performance decisions simultaneously, the following holds for the platform performance at equilibrium:*

i) ϕ^{PS} *increases with end-users' utility from an additional unit of platform performance,* β_c

ii) ϕ^{PS} *decreases with end-users' utility from an additional unit of content available,* α_c

iii) ϕ^{PS} *decreases with royalty rate,* γ.

iv) ϕ^{PS} *decreases with content development cost per unit performance,* β_g.

v) ϕ^{PS} *does not depend on the intensity of competition in the end-user market,* t

Proposition 3.4.3 shows us that in a simultaneous-move game, if the platform sponsors have the degree of freedom to set the end-user price in addition to the platform performance, the equilibrium performance in the market is immune to the counterintuitive effects of two-sidedness that were presented in Proposition 3.4.1. The reason behind this is two-folds. First, note that most of the counterintuitive effects in Proposition 3.4.1 are observed in a content driven-market where $v_\phi = \beta_c - \alpha_c\beta_g \leq 0$. However, when v_ϕ is negative, so is v_{PS} implying that the price-setting platforms choose not to invest in platform performance at all (by Lemma 3.4.2). Note that high performance is costly to the platform providers in two ways, being the fixed cost of developing the platform and the risk of reduced participation from the developer side. Hence, in a content driven market where the value of platform performance is low, the tendency to avoid investing in platform performance is intuitive. In the absence of a second leverage though, price-taker platforms are not always able to avoid investing in platform performance; rather they manage that investment to balance the cross-network effects, whereas price-setting platforms can and do avoid investing. But this implies that performance choice no longer provides leverage for price-setting platforms in a content driven market as their key leverage becomes the end-user price. As a result, the counterintuitive effects observed for price-taker platforms do not apply to price-setting platforms.

Table 3.2: Comparative Statics in the simultaneous-move game

	Conventional Wisdom	*Price Taker Duopoly* (Proposition 3.4.1)	*Price Setting Duopoly* (Proposition 3.4.3)
β_c	$\phi^* \uparrow$	$\phi^* \uparrow or \downarrow$	$\phi^* \uparrow$
t	$\phi^* \uparrow$	$\phi^* \uparrow or \downarrow$	$\phi^* doesn't change$
α_c	$\phi^* \downarrow$	$\phi^* \uparrow or \downarrow$	$\phi^* \downarrow$
γ	$\phi^* \downarrow$	$\phi^* \uparrow or \downarrow$	$\phi^* \downarrow$
α_g	$\phi^* \uparrow$	$\phi^* \uparrow or \downarrow$	$\phi^* \uparrow$

The second reason has to do with the curse of choice. In a competitive setting, the additional pricing power may trigger a price war which constrains the ability of platforms to adopt some counterintuitive yet profitable strategies. For instance, Part iv) of Proposition 3.4.1 suggests that a price-taker platform may choose to increase the platform performance when end-users' utility from content availability, α_c, increases. On the other hand, price-setting platforms take the more intuitive action and reduce the platform performance. However they also cut back the end-user price in order to stay competitive whereas commitment to the platform price enables the price-taker platforms to afford an increase in the platform performance, which turns out to be a more profitable strategy under certain market conditions.

Comparing the profits earned by price-taker versus price-setting platforms reveal the following:

Proposition 3.4.4. *Consider two symmetric platforms simultaneously entering the market.*

i) If end-users' utility from platform performance, β_c, increases, the profit of a price-setting platform always decreases. However, the profit of a price-taker platform actually increases if the market is content driven.

ii) If the degree of competition increases, the profit of a price-setting platform always decreases. However, the profit of a price-taker platform actually increases if the market is content driven.

Proposition 3.4.4 provides another example for how the additional pricing power may be a drawback. For example in the presence of pricing power if competition intensifies, platforms engage in a price war, which draws down the profits. However, as shown in Proposition 3.4.1, price-taker platforms may respond to increasing competition by lowering the platform performance in a content-driven market. In this case, platform development costs go down and the platform becomes more attractive for developers with its low cost of content development. End-users lose some utility due to the reduction in platform performance, yet this loss is easily compensated by increased content availability. Hence, price-taker platforms simultaneously entering the market may actually benefit from increased competition if the market is content driven. Similarly, an increase in end-users' utility from platform performance may benefit price-taker platforms in a content-driven market, but never helps price-setting platforms.

Finally, contrary to the intuition that high performance goes together with high price, we observe that in response to a change in certain market characteristics, end-user price may increase despite a decrease in the platform performance.

Proposition 3.4.5. *Consider a duopoly of price-setting platforms simultaneously entering the market.*

i) When content development cost per unit performance, β_g, increases, equilibrium price p^{PS} may increase even though equilibrium performance ϕ^{PS} always decreases.

ii) When royalty rate, γ, increases, equilibrium price p^{PS} may increase even though equilibrium performance ϕ^{PS} always decreases .

An increase in content development cost per unit performance results in a decrease in platform performance at equilibrium. Accordingly, in the absence of cross-network effects one would expect the equilibrium price to decrease on the grounds that a lower performance platform would deserve a lower price. However, Proposition 3.4.5 shows that optimal price may actually increase because by reducing the platform performance, it is possible to attract more content developers which may justify a higher end-user price. Similarly, end-user price may increase when the royalty rate increases despite an accompanying decrease in platform performance.

3.5 Sequential-move game

In this section, we analyze two competing platforms, one of them being the market leader. The follower enters the market after observing the decisions made by the leader, i.e. the core performance of the leader's platform and if the platforms are not price takers, the end-user price set by the leader. The leader anticipates the follower's response and takes that into account in making his decisions. Note that the sequential move-games analyzed in this paper are not dynamic in nature. We assume that even though the follower enters the market late, the leader does not accumulate an installed base in the mean time. This assumption would fit to a situation where the two platform sponsors make their decisions sequentially but they deliver the products to the market simultaneously. Another possible scenario is that the follower announces its entry in advance and rational agents (end-users and content developers) anticipate the performance and price decisions of the follower. Accordingly, if the agents prefer the follower' s platform, they wait for the release of that platform without incurring a loss of utility due to waiting.

In the sequential-move game even though closed form solutions for the equilibrium exist, the results are quite intractable, especially for the case of price-taker platforms. Hence, we briefly comment on the price-taker duopoly based on numerical experiments and further elaborate on the price-setting duopoly.

3.5.1 A price-taker duopoly

When we solve the decision problem (3.7) sequentially using backwards induction, we first observe that unlike the simultaneous-move game with symmetric platforms, the equilibrium is not necessarily symmetric. Indeed, the follower may make a lower or a higher performance investment than the leader even if they commit to the same end-user price. Whether the follower chooses a higher core performance depends on a complicated set of factors. For example, in Figure 3.4, the dashed line is the follower and the solid line is the leader. The follower commits to a slightly higher end-user price. Yet, committing to a higher price does not necessarily imply investing more on platform performance. When end-users' utility from platform performance is high, the leader adopts a more aggressive strategy of performance investment despite its price advantage.

More importantly, we observe that the counterintuitive comparative statics results obtained in Section 3.4.1.1 still hold qualitatively for the sequential-move game. In other words, irrespective of the type of game played between

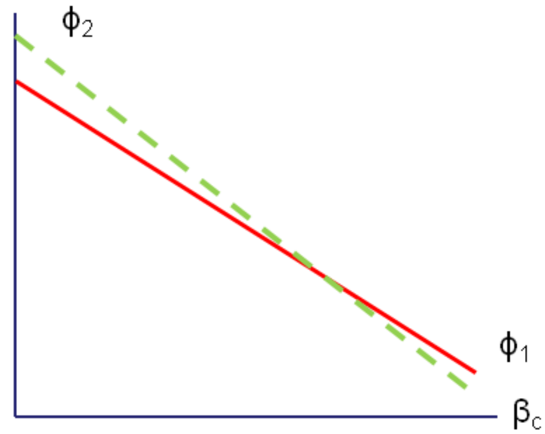


Figure 3.4: The follower's platform may have a higher or a lower performance than the leader's

price-taker platforms, it is possible to have the equilibrium performance decrease with end-users' utility from platform performance β_c, product differentiation in the end-user market, t and content price α_g and increase with end-users' utility from content availability, α_c and the royalty rate, γ.

3.5.2 A duopoly with price-setting platforms

Analyzing a sequential-move game between price-setting platforms uncover interesting insights on the different strategies employed by the market leader and the follower. For instance, the follower and the leader respond differently to changes in the content price α_g and in the degree of product differentiation t, as shown in Proposition 3.5.1.

Proposition 3.5.1. *Consider a price-setting duopoly of platforms. The leader and the follower respond in opposite directions to changes in:*

i) content price α_g

ii) degree of product differentiation t

If the leader increases its platform's performance in response to a change in α_g or t, the follower reduces, and vice versa.

The leader indeed takes advantage of moving first when responding to the changes in content price. For example, in a market dominated by casual gamers, when game price increases it may be more profitable to further reduce the platform's core performance to attract more game developers because ultimately that is what gamers value the most. However, we see that when the leader adopts this strategy, it is no longer profitable for the follower to adopt it. Hence, the follower ends up increasing the performance of its platform. We observe a similar strategy difference in platforms' responses to changes in the degree of competition. Note that in the simultaneous-move game between price-setting platforms, a change in the degree of competition is only reflected in the equilibrium price decision. However, if there is a leader and a follower in the market, competition in the end-user market affects the performance decisions as well. More interestingly, the leader and the follower respond in opposite directions. In particular, we observe that when competition decreases, the leader takes advantage by decreasing the performance of the platform and increasing the end-user price as long as end-users' utility from platform performance is not too high. However, this leaves the follower with little room to decrease its performance in a profitable manner so the follower has to differentiate itself by increasing the platform performance instead.

3.6 Discussion

In platform markets, “winner-take-all” outcomes are commonly observed due to network externalities. Often, platforms engage in fierce competition and the winner becomes a standard, as was the case with Windows, PDF, DVD, fax and others. Accordingly, the conventional wisdom is to heavily invest in the core performance or features of the platform to get the upper hand in the race. However, our results show that high performance does not always gives the competitive edge. In content-driven markets, the platform with a lower performance can indeed become the market leader. Nintendo Wii’s success despite its low performance provides an example for that.

Our model characterizes a content-driven market as one with low *market value of performance* yet with high degree of competition between the platforms. In terms of the end-user preferences, this implies more emphasis on content availability and variety than the core performance of the platform. In terms of the content developers, this implies significantly higher development costs when developing for high performance platforms. As the video-game industry expanded its target market beyond the “hard core gamers” to include more and more “casual gamers”, demand for better performance has been on the decline (Sheffield 2008), triggering a shift to a content-driven market. Adding the ever rising costs of game development for the current generation technology, a platform like Wii that almost relies on the previous generation technology gains a competitive advantage in attracting game developers. In a market where it is the games that matter, once the game developers are on board, so are the gamers.

It took the game industry some time to recognize that the average gamer does not ask for better graphics or better sound capabilities than what Nintendo Wii offers (Sheffield 2008). Indeed, there is usually a temptation to overinvest in the core technology of a platform. The fundamental message this paper delivers is that outperforming the competition sometimes just boils down to facilitating third party development better.

We used a stylized model to uncover these insights. Perhaps the most restrictive of our assumptions is that the end-user market size is fixed. This assumption is implicit in the Hotelling model, variants of which are frequently used in the two-sided markets literature (Parker and Alstyne 2000, Armstrong 2002, Rochet and Tirole 2003, Armstrong 2006, Armstrong and Wright 2007, Anderson and Coate 2005, Kaiser and Wright 2006). Fixed market size simplifies the analysis to a great extent but also presents a limitation because strategies that could expand or fail to attract the total market are not accounted for. Hence, our model is more useful to understand performance investment strategies that sustain the current market rather than those that change the game completely.

Throughout the analysis we assume that content developers are able to multihome whereas end-users choose to join a single platform. This framework fits the video-game industry fairly well; however, it does not directly extend to markets where platforms make exclusivity deals with content developers. If each developer works exclusively for a particular platform, in other words if both sides of the market singlehome, the platform sponsors have to compete in both sides of the market; thus the network effects become more critical. Our preliminary analysis on this framework shows that the main results for the performance investment strategy such as the counterintuitive trends presented in Proposition 3.4.1 continue to hold qualitatively. However, pricing in a two-sided single-homing framework presents some distinctions. A detailed comparison between the two frameworks is left as future research.

Finally, we study single-period models of competition between platforms. It would be interesting to analyze performance investment strategies of incumbents and entrants in a dynamic framework. Our single-period analysis of a sequential move game does not account for the installed base advantage of the early mover, as the market is cleared simultaneously even though the decisions are made sequentially.

3.7 Conclusions

Platform development and design is a dimension of two-sided markets that has not been adequately addressed in the literature. In this paper, we explore the performance investment strategies of competing platforms in a two-sided market. We focus on a platform sponsor's trade-off between developing a high performance platform that matches the end-user's preferences and sacrificing some of those preferences in exchange for improved or less costly third party development capabilities. We show that conventional wisdom about product development decisions may be misleading in the presence of strong cross-network externalities.

We first characterize content-driven and performance-driven markets. Market value of platform performance as well as the degree of differentiation between the platforms is low in the former and high in the latter. Conventional wisdom may be especially misleading in a content-driven market. For instance, in a one-sided market if the degree of competition between the firms is higher, one would expect to see more aggressive investment in the performance of the product. However, we show that in a content-driven market platform sponsors are better off decreasing the investment in platform performance and providing greater content availability instead.

More importantly, we show that when platforms are price-takers, a platform with a lower performance can indeed become the market leader, similar to Nintendo Wii's success against its high performance competitors Xbox and PS3. In other words, contrary to the conventional wisdom about "winner-take-all" markets, heavily investing in the core performance of a platform with strong cross-network effects may not yield a competitive edge. Our results suggest that offering a platform with lower performance but greater availability of content may be a better strategy, particularly in content-driven markets.

When platforms are able to set the end-user price in addition to the platform performance, two-sidedness presents itself in the interplay between end-user price and platform performance. For instance, we show that in response to changes in certain market parameters, it may be optimal to reduce the investment in the performance of the platform and increase the price at the same time, if the reduction in the platform performance is a strong enough lever to attract more content developers. Finally, we show that in a leader-follower game, the follower may have to differentiate its platform from that of the leader by counteracting the leader's performance decisions. For example, when competition decreases, we observe that the leader generally seizes the opportunity by decreasing his investment in platform performance whereas the follower chooses to increase his investment by exactly the same amount.

Chapter 4

Dual Sourcing under Random Supply Capacities: The Role of the Inferior Supplier

4.1 Introduction

Multiple sourcing is common in practice due to various reasons. First, firms may make the best cost–responsiveness balance by keeping a cheap offshore supplier and an expensive onshore supplier. Allon and Van Mieghem (2010) provide such an example for a U.S. manufacturer of wireless transmission components with suppliers in China and Mexico. Second, maintaining a portfolio of suppliers allows for risk sharing. Supply diversification has become a common strategy to hedge against possible shortfalls of the supply streams. Tomlin (2006) recounts two examples that underscore the value of multiple sourcing in the occasion of supply disruptions. When Hurricane Mitch hit Central America in 1998, Dole suffered from the supply disruption in the region, while Chiquita was much less affected because of its larger supplier base. A similar story is reported after the fire at the Philips plant in 2000. Ericsson, heavily relying on Philips, was significantly hurt, while Nokia was able to get around by resorting to its alternative suppliers. Finally, when the output volume of each individual supplier is insufficient, multiple sourcing becomes a must. For instance, Apple has been constrained by the low volume from its primary supplier, LG Display, for iPad LCD panels. In an attempt to alleviate this capacity restriction, Apple contracted with Samsung as an additional supplier in early 2010.

In formulating a multiple sourcing strategy, firms must determine the allocation among different suppliers based on their key characteristics including cost, leadtime, reliability and capacity. Intuitively, if one supplier is *inferior* in all dimensions compared to other suppliers, it seems undesirable to order exclusively from this supplier or to allocate a major portion of demand to this supplier. Several previous studies have hinged on these notions. For example, the presence of a fast and cheap supplier excludes a slow and expensive supplier if both are reliable and uncapacitated (e.g., Fukuda 1964, Allon and Van Mieghem 2010). Also, a cheaper supplier should be selected before a more expensive one with the same leadtime if both are unreliable and uncapacitated (e.g. Dada, Petruzzi and Schwarz 2007, Federgruen and Yang 2008). The question is: Do these notions hold in general when the suppliers can differ from one another in cost, leadtime, reliability and capacity?

To provide the answer, we formulate a multi-period replenishment model. Specifically, the firm under consideration has two procurement sources, a fast supplier of one-period leadtime and a slow supplier of two-period leadtime.[1]

[1]Such leadtime structure is commonly assumed in the literature to ensure analytical tractability (e.g., Whittmore and Saunders 1977, Lawson and Porteus 2000, Sethi, Yan and Zhang 2003).

The procurement cost of the fast supplier can be higher or lower than that of the slow supplier. Each supplier operates a random capacity. The firm decides the order quantities in each period before the supply and demand uncertainties are resolved.

For the resulting model, the optimal policy is characterized by two reorder points, one for each supplier. A positive order is only issued to a supplier if and only if the inventory level is below its reorder point. Interestingly, the reorder point for the slow supplier can be higher than that for the fast supplier even if the slow supplier has a higher cost, a lower reliability and a smaller capacity than the fast one. In other words, it is possible to order *exclusively* from a supplier who is inferior in all dimensions –cost, leadtime, reliability, and capacity.

Such a phenomenon can be driven by either capacity limit or capacity uncertainty of the fast supplier. In particular, we prove that regardless of its cost and capacity distribution, the slow supplier's reorder point exceeds that of the fast supplier when the latter's capacity is low enough or when the latter's no-delivery probability is high enough. These conditions dampens the fast supplier's ability of quick replenishment, leading to a risk of stockout. Such a risk can be effectively mitigated by a slow order placed in advance, even if the slow order is expensive and unreliable.

This intriguing observation suggests us to further examine how the firm should allocate procurement orders between the fast and slow suppliers, who are otherwise identical. Our analysis suggests that the allocation rule depends critically on capacity limit and capacity uncertainty. When the fast supplier has an ample reliable capacity, the slow supplier is only used as a backup in the occasion of severe stockout. In contrast, facing a limited or unreliable fast supply, the firm may rely heavily on the slow supplier. Particularly, the slow supplier may obtain more than 50% of the order allocation in the long run even if it has the same cost and capacity distribution as the fast one.

Furthermore, a larger allocation is given to the slow supplier and a smaller allocation to the fast one when the planning horizon is longer. Such an allocation scheme leads to a near-perfect match between supply and demand in the long run—the average delivery quantity approaches the average demand as the planning horizon extends. Consequently, the total order from both suppliers is smaller in an earlier period than in a later period. This observation underscores a key difference between dual sourcing and single sourcing. For the latter, it is well known that less should be ordered when getting closer to the end of the horizon (see, e.g., Federgruen and Zipkin 1986).

We also demonstrate that the effect of capacity limit and capacity uncertainty can play out differently with respect to changes in model parameters. For example, under unlimited and random supply, an increased demand variability leads to an increased allocation to the fast order. This is because the fast supplier, without a capacity limit, is more responsive in reacting to demand shocks than the slow supplier. However, under limited and deterministic supply, a slow order placed in advance can help to avoid stockouts induced by large demand variability when the fast supply capacity is limited. Thus, the allocation to the slow supplier increases in response to an increased demand variability.

Our study highlights the role of an inferior supplier in the context of multiple sourcing. In particular, when the supply capacity is very restrictive or highly uncertain, firms should order primarily from the slow supplier and use the fast order as a supplement, even if the slow supplier does not beat the fast supplier in cost or reliability. This counterintuitive observation underscores the importance of incorporating cost, leadtime, reliability, and capacity in a unified framework to evaluate supplier selection and order allocation strategies.

The remainder of the paper is organized as follows. In §4.2, we review the related literature and discuss our contributions. The model is laid out in §4.3 and the optimal policy is characterized in 4.4. In §4.5, we analyze order allocation between the fast and slow suppliers via an extensive numerical study. Section 4.6 concludes and points out the future research directions. All the proofs are relegated to Appendix C.

4.2 Literature Review

This paper brings together two streams of research. The first concerns sourcing from multiple reliable suppliers, each characterized by a delivery leadtime and a procurement cost. This line of study dates back to Barankin (1961), Daniel (1963) and Fukuda (1964) with further developments by many others. It is commonly assumed that a supplier with a longer leadtime charges a lower procurement cost than one with a shorter leadtime. In general, the problem involves a large state space of in-transit orders due to the possibility of order crossing, i.e., an order placed earlier may arrive later. When this happens, the optimal policy can be quite complex (Whittmore and Saunders 1977, Feng, Sethi, Yan and Zhang 2006) and one has to resort to heuristics (e.g., Scheller-Wolf, Veeraraghavan and Houtum 2005, Veeraraghavan and Scheller-Wolf 2008, Sheopuri, Janakiraman and Seshadri 2010). For periodic-review systems, one way to exclude order crossing is to assume two suppliers with consecutive delivery leadtimes, i.e., the leadtimes vary by exactly one period. In fact, the main analytical development of this literature is restricted to this case (e.g., Fukuda 1964, Lawson and Porteus 2000, Sethi et al. 2003, Yazlali and Erhun 2009), where the optimality of a base-stock policy is established. We also assume two suppliers with consecutive leadtimes for analytical tractability. However, our model adds another layer of complexity by incorporating uncertain supply capacities, which leads to the suboptimality of a base-stock type policy.

For the problem of multiple sourcing from reliable suppliers, there are other ways to reduce the state space of in-transit orders. First, one can restrict ordering at certain periods and thus eliminate order crossing. Allowing orders to be placed every N periods, Fukuda (1964) shows that a base-stock policy can be obtained when the difference between the longest and shortest supply leadtimes is N periods. Under dual sourcing, this assumption implies at most one outstanding order at any time (e.g., Chiang and Benton 1994). Second, one can consider a model with a standing order from a slow supplier and dynamic orders from a fast supplier (Allon and Van Mieghem 2010). There has not been any consideration of supply capacity uncertainty in these models. Our analysis suggests that the structure of the optimal policy characterized for our model carries through in the first case. For the second case, the dynamic order from the fast supplier will follow a base-stock policy as suggested by Ciarallo, Akella and Morton (1994). In both cases, the main insight obtained from our model will be preserved. That is, an advance order from the slow supplier provides the advantage of mitigating the stockout risk. Therefore, the firm may rely heavily on a slow supplier even if it does not offer cost or reliability benefit over the fast supplier.

It is worth contrasting our result to that obtained by Allon and Van Mieghem (2010). They propose a tailored base-surge allocation strategy, which assigns the base demand to a cheap offshore supplier and the surge demand to an expensive onshore supplier. This strategy effectively captures the cost–flexibility trade-off between a slow supplier and a fast supplier. Because their model does not impose a capacity limit for fast supplier, the allocation policy derived in their model breaks down if the slow supplier is more expensive than the fast one. Our model, however, allows such a possibility. While our analysis also suggests using the slow supplier as the primary source and the fast one as a supplement, the reason differs from theirs. In their case, the slow supplier obtains a large allocation due to its advantage of low procurement cost, whereas such a situation can arise in our model even if the slow supplier is more expensive.

The second stream of related research examines supply uncertainty (see the survey by Yano and Lee 1995), which is commonly modeled as random yield (Henig and Gerchak 1990), random capacity (Ciarallo et al. 1994, Yang, Qi and Xia 2005, Chao, Chen and Zheng 2008), or a combination of both (Wang and Gerchak 1996). Different strategies have been proposed for supply risk mitigation. For example, Kouvelis and Li (2010) examine the value of contingent response to observed yield uncertainty. Wang, Gilland and Tomlin (2010) compare the benefit from improving supply reliability with that from dual sourcing. Many studies also underscore the importance of supply diversification in risk mitigation. With continuous demand distributions, it is optimal to select a cheaper supplier before a more expensive one regardless of the suppliers' reliability (Anupindi and Akella 1993, Dada et al. 2007, Burke, Carrillo and Vakharia 2007, Federgruen and Yang 2008, Federgruen and Yang 2009). With discrete demand distributions, Swaminathan and Shanthikumar (1999) demonstrate the possibility of ordering only from the expensive supplier whose reliability is significantly higher than the cheaper supplier. All of these studies overlook the difference of suppliers' leadtimes.

Our study combines these streams of literature by examining cost, leadtime, reliability and capacity in a unified framework. We demonstrate the possibility of ordering exclusively from and relying heavily on an inferior supplier. These observations are in stark contrast to the earlier findings.

4.3 The Model

The firm faces a T-period planning problem. At the beginning of period t, the inventory level I_t is reviewed. Then, the firm makes replenishment decisions for an uncertain demand D with the options to source from a fast supplier and a slow supplier. In particular, the fast supplier has a delivery leadtime of one period and the slow one has a leadtime of two periods. When a fast order of $q_{f,t}$ and a slow order of $q_{s,t}$ are issued, the amounts delivered are random due to the uncertainties associated with the supply capacities k_f and k_s, respectively. We assume that k_f, k_s, and D are stationary and independent. The uncertain capacities k_f and k_s represent the random resource availability levels for providing the product. For example, the machines and equipment deployed for production may break down from time to time or deteriorate over time. The productivity level may fluctuate due to variations in worker skills or operating conditions. In the case of external sourcing, there may be a lack of information about the suppliers' production capabilities and workforce levels. Even if the information is shared, the suppliers are typically reluctant to reveal other buyers' orders, which often have a major influence on the available supply capacities of the system under consideration.

The ordering quantities from both suppliers must be specified before the uncertainties in the demand and supply are resolved. At the end of period t, k_f, k_s and D are observed. The unmet demand is backordered and the leftover inventory is carried over to the next period. Specifically, the ending inventory for period t is

$$I_{t+1} = I_t + q_{f,t} \wedge k_f + q_{s,t-1} \wedge k_s - D, \tag{4.1}$$

where $a \wedge b = \min\{a, b\}$.

The firm pays c_j dollars for each unit received from the order $q_{j,t}$, $j = f, s$. He also pays a surplus/shortage cost $H(\cdot)$ upon the demand realization, defined as

$$H(x) = hx^+ + px^-, \tag{4.2}$$

where $x^+ = \max\{0, x\}$ and $x^- = \max\{0, -x\}$. We assume $c_f < p$ and $c_s < \alpha p$ to rule out the trivial cases of never ordering from the fast supplier and the slow supplier, respectively.

Let $V_t(I)$ denote the optimal cost function in period t when the inventory level is I. Then, the dynamic programming equation is given by

$$V_t(I) = \min_{q_{f,t} \geq 0, q_{s,t} \geq 0} J_t(I; q_{f,t}, q_{s,t}),$$

where

$$\begin{aligned} J_t(I; q_{f,t}, q_{s,t}) &= c_f \mathbb{E}(q_{f,t} \wedge k_f) + c_s \mathbb{E}(q_{s,t} \wedge k_s) + \mathbb{E}H(I + q_{f,t} \wedge k_f - D) \\ &\quad + \alpha \mathbb{E} V_{t+1}(I + q_{f,t} \wedge k_f + q_{s,t} \wedge k_s - D). \end{aligned}$$

Finally, we impose, without loss of generality, a terminal condition $V_{T+1}(I) = 0$; yet it should be noted that our analysis also extends to the case of infinite horizon.

We shall remark that the policy structure derived from our model carries through for systems with nonstationary cost parameters and those with Markov-modulated supply and demand processes. Our analysis also extends to the case when the leadtimes for the fast and slow suppliers are L and $(L+1)$ periods, provided that the capacities are observed one period after ordering (see Remark C.0.1).

Using standard dynamic programming arguments, we can show that an optimal feedback solution always exists (see Theorem C.0.1 in the appendix). We denote $q^*_{f,t}(I)$ and $q^*_{s,t}(I)$ to be the optimal fast and slow orders, respectively, in period t. When multiple solutions exist, we pick the ones with the smallest slow order and then the one with the smallest fast order.

4.4 The Analysis

In this section, we present a detailed analysis of the problem formulated in §4.3. The properties of the dynamic programming equation obtained in§4.4.1 facilitate further analytical derivations of the optimal policy in §4.4.2.

4.4.1 The Convexity of the Optimal Cost Function and its Implications

The major difficulty with our model is that the presence of uncertain capacity leads to a *nonconvex* objective function. In particular, $J_t(I; q_{f,t}, q_{s,t})$ is not convex in either $q_{f,t}$ or $q_{s,t}$. We need to understand the behavior of the optimal cost function before exploring the optimal policy analytically.

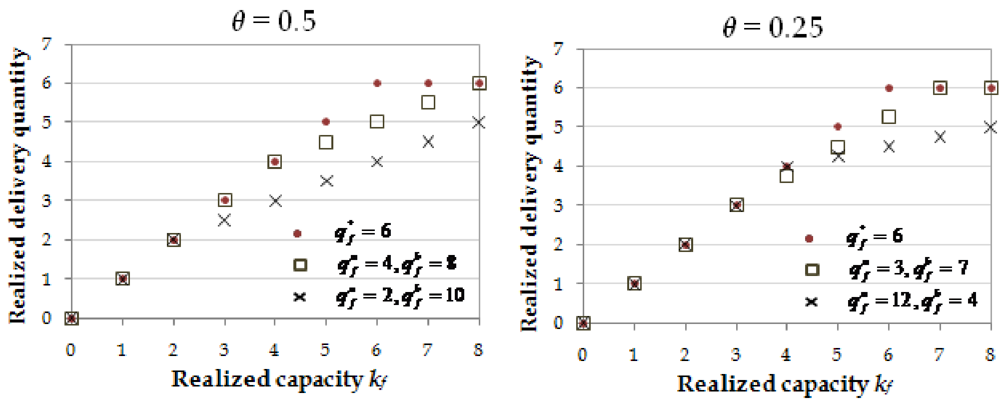
Proposition 4.4.1. *The optimal cost function $V_t(I)$ is convex in I for each t.*

When both the fast and slow orders are optimally determined, the resulting expected optimal cost function $V_t(I)$ is convex in the inventory level I. In other words, the marginal cost of on-hand inventory is increasing. This result generalizes its counterpart in Ciarallo et al. (1994), who consider ordering from a single supplier with random capacity. They derive an optimal base-stock policy, which in turn leads to the convexity of the optimal cost function in their model. In contrast, Proposition 4.4.1 is established without much knowledge of the optimal policy. Therefore, the approach adopted here is very different from theirs. Moreover, as we will see in §4.4.2, the base-stock type policy fails to achieve optimality in our model in general.

Proposition 4.4.1 is established based on the following observation (see a formal treatment in Lemma C.0.2 in Appendix). For a fixed order quantity from, say, the slow supplier, we compare the expected cost under two alternatives. The first is to order the optimal quantity $q^*_{f,t}$ from a single fast source with a capacity of k_f, and obtain $q^*_{f,t} \wedge k_f$ units. In the second alternative, there are two fast suppliers whose capacities, denoted by k^a_f and k^b_f, are perfectly positively correlated and have the same distribution as k_f. When orders q^a and q^b are issued, the delivery quantities are $\theta(q^a \wedge k^a_f)$ and $(1-\theta)(q^b \wedge k^b_f)$, for some arbitrarily given $\theta \in (0, 1)$. It turns out that the first alternative always dominates the second. A similar result can be obtained when we switch the fast and slow suppliers in the above arguments. Taken together, we reach the conclusion of Proposition 4.4.1.

The above observations lead to some interesting insight—*while diversifying among independent suppliers leads to a cost saving, split orders among positively correlated sources may dampen system performance.* This result shares a similar spirit with the well-known observation from the risk pooling literature on the negative effect of pooling positively correlated demands. This is because the variance of the pooled demand is higher than the sum of the individual demand variances. In the second alternative mentioned above, in contrast, when splitting the orders between two sources with perfect positive correlation, the distribution of the total supply capacity remains unchanged. Still, ordering from two sources underperforms ordering from one.

To see the intuition, suppose $q^*_{f,t} = 6$ for some fixed I and $q_{s,t}(I)$. Then by Ciarallo et al. (1994), $q^*_{f,t} = 6$ is also the optimal order quantity when the fast supply is unlimited and perfectly reliable. Now consider a random capacity k_f uniformly distributed over $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$. We demonstrate the delivery quantity for each realized capacity k_f under different ordering alternatives in Figure 4.1. The benchmark (the dots) is to order $q^*_{f,t}$ and obtain $q^*_{f,t} \wedge k_f$. The squares and crosses correspond to different way of splitting $q^*_{f,t}$ by (q^a, q^b) such that $\theta q^a + (1-\theta)q^b = q^*_{f,t}$. We first observe that in all cases, the realized delivery quantity is no higher than the desired $q^*_{f,t}$. Furthermore, the benchmark case always has a higher realized delivery quantity, no matter how we split $q^*_{f,t}$ between q^a and q^b. Therefore, the benchmark case yields the minimum expected cost.



Notes. The capacity k_f is uniformly distributed over $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$. The first alternative (represented by the dots) orders $q_f^* = 6$ from a source of capacity k_f. The second alternative (represented by the squares or the crosses) orders q^a and q^b with $\theta q^a + (1-\theta)q^b = q_{f,t}^*$ and obtains $\theta(q^a \wedge k_f) + (1 - \theta(q^b \wedge k_f)$.

Figure 4.1: An example of the delivery quantity with respect to realized capacity level.

4.4.2 The Optimal Policy

The convexity of the optimal cost function established in the last section allows us to further characterize the optimal policy. For ease of exposition, we assume that the demand has a continuous density so that the optimal cost function $V_t(I)$ is differentiable. This is commonly assumed in the literature (e.g., Fukuda 1964, Anupindi and Akella 1993, Ciarallo et al. 1994, Dada et al. 2007). It is clear, though, from our subsequent analysis that our results can be extended to the case of discrete demand.

When the suppliers are perfectly reliable (Fukuda 1964, Yazlali and Erhun 2009) or when there is only one supplier with random capacity (Ciarallo et al. 1994), the optimal policy is known to be a base-stock type. However, when purchasing from two sources with random capacities, a base-stock policy is generally suboptimal, as indicated in the next proposition.

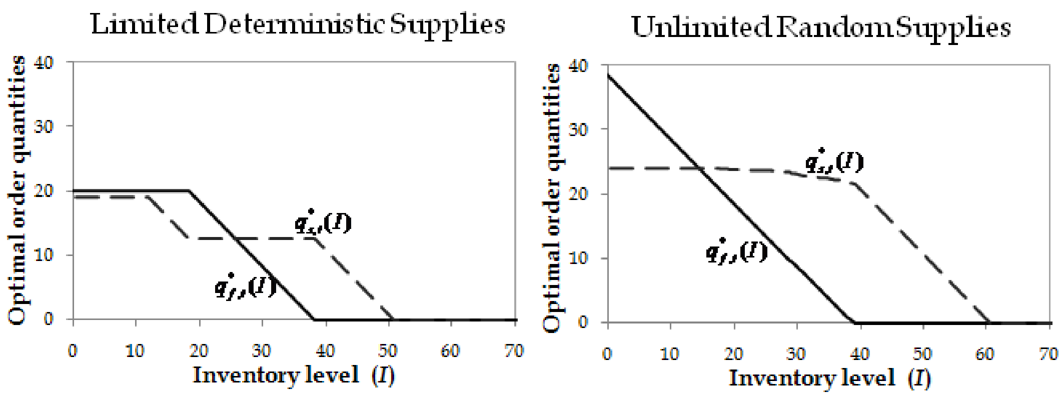
Proposition 4.4.2. *Under the optimal ordering policy, the following relations hold for a small enough positive δ.*

$$0 \le q^*_{f,t}(I-\delta) - q^*_{f,t}(I) \le \delta \textit{ for } \forall I, \tag{4.3}$$

$$0 \le q^*_{s,t}(I-\delta) - q^*_{s,t}(I) \le \delta \textit{ for } \forall I, \tag{4.4}$$

$$[q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta)] - [q^*_{f,t}(I) + q^*_{s,t}(I)] \ge \delta$$
$$\textit{for } I \le \max\{I_{f,t}, I_{s,t}\}. \tag{4.5}$$

When the inventory level decreases, the fast and slow orders should increase. This suggests that there exists a reorder point I_j, $j = s, f$, such that a positive order $q^*_{j,t}(I)$ is issued if and only if I is below I_j. When the suppliers are both reliable (Fukuda 1964), I_j becomes the base-stock level (see the left panel of Figure 4.2). In this case, for $I \le \max\{I_f, I_s\}$, a unit reduction in inventory implies exactly one unit increase in one of the procurement orders. Similarly, when only one supplier with random capacity k_j is involved (Ciarallo et al. 1994), a unit reduction in inventory is always compensated by a unit increase in $q^*_j(I)$ for $I \le I_j$. In the general case of two suppliers with random capacities, though, each procurement order compensates a unit decrease in inventory by an amount no more than one unit, as stated in Proposition 4.4.2 and demonstrated in the right panel of Figure 4.2. Furthermore, the total order quantity increases by at least one unit, indicating a tendency to mitigate the delivery uncertainty. In other words, the optimal policy not only splits the bet between the two suppliers to share the risk, but also increases the total bet to ensure a certain post-delivery stock level.



Parameters: $T = 10$, $\alpha = 0.95$, $c_f = 9$, $c_S = 10$, $h = 2$, $p = 21$, and $D \sim$ Normal$(30, 6)$. The left panel (limited deterministic supplies) assumes $k_f = 20$ and $k_s = 19$. The right panel (unlimited random supplies) assumes $k_f, k_s \in \{0, \infty\}$ with $\Pr\{k_f = 0\} = 0.15$ and $\Pr\{k_s = 0\} = 0.2$.

Figure 4.2: The optimal ordering quantities as functions of the inventory level.

In general, the optimal reorder point for the fast supplier can be higher or lower than that for the slow supplier, depending on the cost structure. This is summarized in the next proposition.

Proposition 4.4.3. *The following relations hold for the reorder points*

i) *If* $c_s > h + c_f$, *then* $I_{s,t} \leq I_{f,t}$.

ii) *If* $c_f > c_s + p$, *then* $I_{f,t} \leq I_{s,t}$.

The condition $c_s > h + c_f$ indicates that buying one unit from the slow supplier is more expensive than buying one unit from the fast supplier and carrying that unit for one period. In this case, the reorder point for the slow supplier is always lower than that of the fast supplier. In other words, it is suboptimal to order exclusively from the slow supplier, as Proposition 4.4.3 i) suggests.

When $c_f > c_s + p$, buying one unit from the fast supplier is more expensive than backordering that unit for one period and procuring it from the slow supplier. In this case, the reorder point for the fast supplier is lower than that of the slow. Thus, the optimal policy never calls for ordering exclusively from the fast supplier, as stated in Proposition 4.4.3 ii).

Proposition 4.4.3 generalizes an earlier result derived for the special case of unlimited and reliable supplies (Fukuda 1964), i.e., $k_f = k_s = \infty$. The condition in i) implies $c_s \geq c_f$, under which it is obviously suboptimal to ever order from the slow supplier in this special case. If, however, $c_s < \alpha c_f$, then it is known that the reorder point for the slow supplier is always above that of the fast supplier. In other words, with enough inventory for the current period, there is no need to order from the expensive fast supplier. Yet, the low-cost slow supplier can be used to build up inventory for the subsequent periods.

4.4.3 The Possibility of Ordering Exclusively from an Inferior Supplier

An important feature of the model with $k_f = k_s = \infty$ is that a supplier with inferior characteristics in both leadtime and cost should be completely excluded. A similar conclusion has been reached in analyzing the models with multiple unreliable suppliers. Assuming that every supplier has one-period leadtime, it is never optimal to order exclusively from a supplier who has a higher cost and a lower reliability than others (see, e.g., Dada et al. 2007, Swaminathan and Shanthikumar 1999). All of these observations suggest that if only one supplier order is issued, then that supplier must be superior in some way. The question is: Does this notion hold in general?

Interestingly, the answer is no. *It can be optimal to order* **exclusively** *from a supplier who has a higher cost, longer leadtime and lower reliability than the other supplier.* This can be a consequence of two factors as demonstrated by the two examples in Figure 4.2.

The first factor is the capacity limitation of the fast supplier. The example in the left panel assumes reliable supplies with deterministic capacities. We observe that only a slow order is issued for $I \in [38.2, 50.7]$, even though the slow supplier is more expensive. This is because the on-hand inventory is high enough to avoid a fast order in the current period. However, because of the limited fast supply, ordering some units from the slow supplier can help to avoid shortage in the future periods. In this case, it is optimal to order exclusively from the slow supplier who has a longer leadtime and higher procurement cost than the fast supplier.

The second factor that may lead to a similar observation is uncertainty associated with the fast supplier. The example in the right panel of Figure 4.2 assumes unlimited random supplies. For $I \in [39.1, 60.6]$, it is optimal to source only from the slow supplier who is more expensive and less reliable than the fast one. This is a consequence of the supply unreliability, which can be mitigated by splitting the bet between the fast and slow suppliers. Specifically, the slow order issued in the current period shares the risk with the fast order in the next period to meet the demand in that period.

To sum up, the optimal policy may call for ordering exclusively from an inferior supplier to alleviate capacity limitation and to share supply risk. These observations highlight the importance of incorporating cost, leadtime and reliability in a unified framework to evaluate supplier selection strategies.

4.5 Order Allocation between the Suppliers: A Numerical Analysis

Applying the optimal policy derived in the previous section, we further examine how the firm allocates the orders between the fast and slow suppliers via a numerical study. The previous studies on multiple sourcing mainly focus on deriving the optimal ordering policy. The long-term order allocation among suppliers has been mentioned in only a few studies. Yazlali and Erhun (2009) analyze a version of our model with deterministic upper and lower order limits. Federgruen and Yang (2009) suggest an allocation rule according to shortfall probability in their one-period model of multiple suppliers with random yields and one-period leadtimes. Allon and Van Mieghem (2010) derive a formula governing strategic allocation between an expensive nearshore source and a cheaper offshore source under a tailored base-surge policy. Our subsequent analysis of long-term order allocation, in contrast, combines the main supply characteristics including cost, capacity, reliability, and leadtime, which leads to new insights.

For the base case, we consider a planning horizon of $T = 50$ with a discount factor of $\alpha = 0.95$ and an initial inventory of $I_1 = 0$. The firm pays $c_f = c_s = 3$ for each unit obtained from the fast or the slow supplier. The inventory cost function is given by $H(x) = hx^+ + px^-$ with $h = 2$ and $p = 12$. The demand follows a truncated normal distribution with mean $\mu_D = 30$ and $\sigma_D = 9$.

In our subsequent analysis, we focus on the order allocation between the suppliers. Let $\bar{q}_f = \sum_{t=1}^{T} \mathbb{E}q^*_{f,t}(I_t)/T$ and $\bar{q}_s = \sum_{t=1}^{T-1} \mathbb{E}q^*_{s,t}(I_t)/(T-1)$ be the average order per period from the fast and slow suppliers, respectively. For ease of exposition, we simply term $\bar{q}_f$ and $\bar{q}_s$ as the *per-period fast and slow orders*, respectively, with the understanding that they are expectations over all the sample paths of the optimal orders from period 1 to T. We compute the percentage shares of the fast and slow suppliers in the total order as

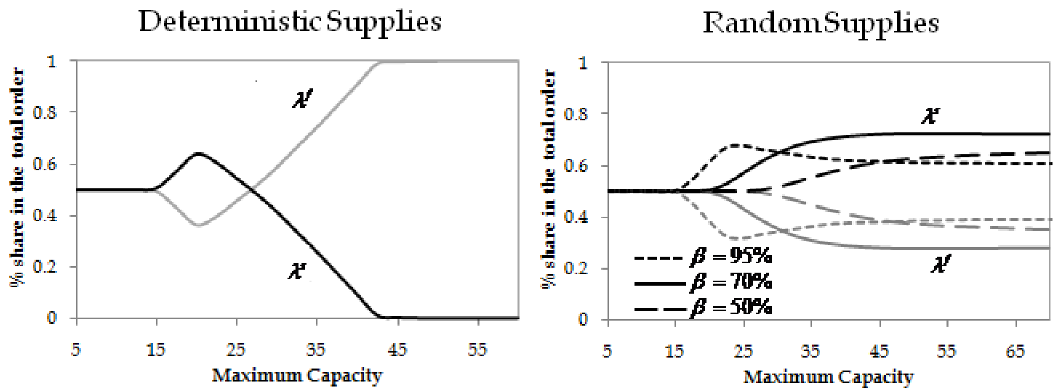
$$\lambda^f = \frac{\bar{q}_f}{\bar{q}_f + \bar{q}_s} \times 100\% \quad \text{and} \quad \lambda^s = \frac{\bar{q}_s}{\bar{q}_f + \bar{q}_s} \times 100\%.$$

Because capacity is an important factor in our model, we first examine its effect on λ^f and λ^s. In the left panel of Figure 4.3, we assume that the suppliers have the same *deterministic* capacity $\overline{k}$. With scarce supply, i.e., $\overline{k} < 13$, the capacities of both suppliers are exhausted, leading to a 50–50 split between them, i.e., $\lambda^f = \lambda^s = 50\%$. For intermediate capacity levels, i.e.,

$\overline{k} \in [13, 26]$, the slow supplier obtains an increased allocation as $\overline{k}$ increases. This is because the fast order, though being reactive, would not meet the demand alone. Given this restriction, it is inevitable to rely on an advance slow order in order to hedge against the stockout risk, which costs much more than the overstock risk as p is higher than h. Consequently, the allocation to the slow supplier λ^s is above 50% in this range of $\overline{k}$. However, the situation is reversed when there is an ample fast supply capacity. Now the need for a slow order is much reduced because the fast order can quickly bring the stock to the desired level. In the extreme case when $\overline{k} > 50$, the share of the slow order λ^s becomes negligible.

The right panel of Figure 4.3 depicts three examples of *random* supply. Specifically, each supplier has a capacity of $\overline{k}$ with probability β and zero with probability $1 - \beta$, where $\beta \in \{0.5, 0.7, 0.95\}$. To our surprise, the allocation to the fast order λ^f decreases in the maximum capacity $\overline{k}$, while that of the slow order λ^s increases. For example, when $\beta = 0.7$, in the extreme case where $\overline{k} \to \infty$, $\lambda^f = 28\%$ and $\lambda^s = 72\%$. Even with a high probability of delivery ($\beta = 0.95$), λ^f is only 40 % when $\overline{k} \to \infty$. *Why should the firm heavily rely on the slow supplier who is less responsive than the fast supplier, yet has the same cost and capacity as the latter?* The reason is twofold. First, because of supply uncertainties, it is beneficial to split the order between the two suppliers to share the delivery risk. Second, when determining the fast order for a given period, the delivery quantity from the advance slow order is already observed. A large fast order is only needed if there is no delivery from the slow order.

As a result, the fast order is only used as a supplement for the slow order, thereby leading to a small λ^f and a large λ^s.



Parameters: $T = 50$, $\alpha = 0.95$, $c_f = 3$, $c_s = 3$, $h = 2$, $p = 12$, and $D \sim$ Normal$(30, 9)$. The left panel assumes limited deterministic supplies. The right panel assumes limited random supplies with $k_j \in \{0, \overline{k}\}$ where $\Pr\{k_j = \overline{k}\} = \beta$, $j = f, s$.

Figure 4.3: The percentage shares of the fast and slow suppliers with respect to the maximum capacity limit.

From Figure 4.3, we observe that the effects of supply limit and supply uncertainty on order allocation can be very different. In order to isolate these effects, we analyze the following capacity configurations.

- *Limited deterministic supply.* In this case, the capacity of either supplier is $\overline{k} = 35$.

- *Unlimited random supply.* In this case, the capacity of either supplier is either infinity or zero with probabilities $\beta = 0.7$ or $1 - \beta = 0.3$, respectively. Thus, the average capacity is infinite.

- *Limited random supply.* In this case, the capacity of either supplier is either $\overline{k} = 35$ or zero with probabilities $\beta = 0.7$ or $1 - \beta = 0.3$, respectively. We shall note that the qualitative insights do not change when we apply other distributions with finite mean, e.g., normal, uniform, etc.

Next we examine the optimal order allocation with respect to model inputs.

4.5.1 Effect of Planning Horizon

It is worth noting from Table 4.1 that $\bar{q}^f + \bar{q}^s$ is decreasing in the planning horizon T in all scenarios. In other words, on average, more units are ordered when it is closer to the end of the horizon. This makes an interesting contrast to a well-known result for replenishment systems following (modified) base-stock policies (e.g., Heyman and Sobel 1984, Federgruen and Zipkin 1986). In particular, a lower base-stock level is maintained for a later period because a unit ordered earlier has a better chance to be sold. The difference between their observations and ours lies in the fact that our model considers two suppliers, as opposed to a single supplier. To understand the intuition, we examine the examples in Table 4.1.

An early slow order has an advantage of hedging the stockout risks induced by the capacity restriction on the late fast order, but has a disadvantage of increasing the risk of overstock. When there are more periods to go, overstock is less of a concern given the many opportunities to sell the units before the end of the horizon. This leads to an increased per-period slow order $\bar{q}^s$ as demonstrated in Table 4.1. With a large $\bar{q}^s$, the available stock at the beginning of a period is likely to meet the demand in that period, reducing the

Table 4.1: The effect of planning horizon on order allocation

	Limited Deterministic Capacity				Unlimited Random Capacity				Limited Random Capacity			
	$\bar{q}_f$	$\bar{q}_s$	$\bar{q}_f + \bar{q}_s$	λ^f	$\bar{q}_f$	$\bar{q}_s$	$\beta(\bar{q}_f + \bar{q}_s)$	λ^f	$\bar{q}_f$	$\bar{q}_s$	$\beta(\bar{q}_f + \bar{q}_s)$	λ^f
$T = 5$	24.72	7.30	32.02	77.20 %	18.61	30.71	34.52	37.73 %	19.28	29.12	33.88	39.83 %
10	23.40	7.66	31.06	75.34 %	15.09	31.03	32.28	32.72 %	16.21	29.57	32.04	35.40 %
30	22.59	7.77	30.36	74.41 %	12.76	31.19	30.76	29.03 %	13.89	29.96	30.69	31.68 %
50	22.43	7.79	30.21	74.23 %	12.29	31.22	30.46	28.25 %	13.42	30.03	30.42	30.88 %
100	22.30	7.80	30.11	74.09 %	11.94	31.24	30.23	27.65 %	13.06	30.09	30.21	30.27 %

need for a large reactive fast order. Consequently, the per-period fast order $\bar{q}^f$ decreases in the length of the planning horizon.

Furthermore, placing a slow order not only reduces the need for a fast order in the next period, but also decreases the fast order quantity in the current period. This is because possible stockouts at the end of the current period may be covered by the incoming slow order in the next period. Therefore, the increase in the per-period slow order is less than the decrease in the per-period fast order, resulting in a reduced total per-period order $\bar{q}^f + \bar{q}^s$. Under deterministic capacities, the total per-period order becomes closer to the average demand $\mu_D = 30$ as the planning horizon extends. This suggests that by relying more on the slow supplier, the system can achieve *a near perfect match between supply and demand* in the long term. With random supplies, the per-period total order $\bar{q}^f + \bar{q}^s$ is generally higher than that under deterministic capacities due to the possibility of no delivery. The expected per-period order quantity $\beta(\bar{q}^f + \bar{q}^s)$, however, also approaches to μ_D as the planning horizon becomes longer.

From the upper panel of Table 4.2, we observe that as the fast order becomes more expensive, its share in the total order, λ^f, is reduced in all cases. This is intuitive and consistent with the observations from Yazlali and Erhun (2009).

Table 4.2: The effect of procurement and penalty costs on order allocation.

	Limited Deterministic Capacity			Unlimited Random Capacity			Limited Random Capacity		
	$\bar{q}_f$	$\bar{q}_s$	λ^f	$\bar{q}_f$	$\bar{q}_s$	λ^f	$\bar{q}_f$	$\bar{q}_s$	λ^f
$c_f = 1$	29.32	0.81	97.30 %	25.63	17.60	59.28 %	21.86	21.43	50.50 %
2	28.82	1.29	95.72 %	16.63	26.79	38.29 %	16.53	26.86	38.10 %
3	22.43	7.79	74.23 %	12.29	31.22	28.25 %	13.42	30.03	30.88 %
4	4.66	25.93	15.23 %	9.88	33.69	22.67 %	12.11	31.37	27.86 %
5	2.44	28.20	7.95 %	8.43	35.16	19.34 %	11.45	32.04	26.33 %
$p =$ **12**	22.43	7.79	74.23 %	12.29	31.22	28.25 %	13.42	30.03	30.88 %
16	22.15	8.11	73.19 %	11.77	31.83	27.00 %	13.55	29.97	31.14 %
20	21.61	8.69	71.32 %	11.54	32.12	26.43 %	13.79	29.79	31.63 %
24	21.51	8.81	70.94 %	11.48	32.22	26.27 %	14.00	29.62	32.10 %
28	21.24	9.11	69.98 %	11.56	32.18	26.43 %	14.15	29.50	32.43 %

The lower panel of Table 4.2 demonstrates the effect of increasing shortage cost p on order allocation (that of decreasing h is similar). A common observation is that $\bar{q}^f + \bar{q}^s$ is increasing in p. This is consistent with our intuition that large orders should be placed when stockout is costly. Interestingly, though, the individual per-period orders $\bar{q}^f$ and $\bar{q}^s$ may increase or decrease in different scenarios.

In the example of deterministic capacity, the per-period fast order $\bar{q}^f$ decreases and the per-period slow order $\bar{q}^s$ increases as p increases. This is simply because as shortage becomes more expensive, the optimal policy tends to prepare for stockouts via placing a slow order in advance.

Under unlimited random capacity, a similar observation is obtained for small enough p. However, as p becomes very large, the optimal policy tends to increase the allocation to the fast supplier and decrease that to the slow supplier. To see the intuition, we note that holding inventory is now much cheaper than backordering demand. It becomes economical to order early and hold inventory in preparation for future demands in view of the uncertainties associated with both fast and slow suppliers. Therefore, the optimal policy tends to place a large fast order that not only aims at meeting the demand in the current period but also attempts to share the risk of orders to be delivered in future periods. Such an effect becomes stronger under limited and random capacity.

4.5.2 Effect of Demand Uncertainty

Intuitively, as the demand variability increases, the fast order becomes increasingly useful, because it is more responsive than the slow order in adjusting the inventory level based on demand realizations. This is in line with the observations by Allon and Van Mieghem (2010) and Yazlali and Erhun (2009). It is also the case in our model with unlimited and random supply, as shown in Table 4.3.

A contrasting observation is obtained under limited deterministic supply—an increased standard deviation of the demand σ_D leads to a reduced share of the fast order, λ^f. In particular, when the demand is nearly deterministic, i.e., $\sigma_D = 1$, the fast supply is enough to meet the demand since $\overline{k} = 35 > 30 = \mu_D$. As demand gets more volatile, however, a slow order may be necessary in the event of a large demand realization, resulting in a reduced λ_f.

Table 4.3: The effect of demand uncertainty on order allocation.

	Limited Deterministic Capacity			Unlimited Random Capacity			Limited Random Capacity		
	$\bar{q}_f$	$\bar{q}_s$	λ^f	$\bar{q}_f$	$\bar{q}_s$	λ^f	$\bar{q}_f$	$\bar{q}_s$	λ^f
$\sigma_D = 1$	30.01	0.00	100.00 %	9.59	33.87	22.06 %	13.74	29.58	31.72 %
3	29.99	0.03	99.91 %	10.22	33.25	23.51 %	12.73	30.65	29.34 %
5	27.93	2.14	92.87 %	10.87	32.62	25.00 %	12.71	30.70	29.28 %
7	25.13	5.01	83.37 %	11.56	31.94	26.58 %	13.04	30.40	30.02 %
9	22.43	7.79	74.23 %	12.29	31.22	28.25 %	13.42	30.03	30.88 %

Under limited and random capacity, the optimal policy reveals a mixed pattern as suggested from Table 4.3. A high demand variability implies an increased chance of large demand realizations, which in turn induces the need for larger orders. Because capacities are limited and random, the fast and slow orders play different roles in mitigating the demand uncertainty. On the one hand, being placed in advance, one can put a large bet on the slow order. Even if it is not delivered, there is still a chance to replenish by ordering from the fast supplier. On the other hand, with a shorter leadtime, the fast order is more reactive to immediate stockouts induced by demand uncertainty. Taken together, the allocation between the fast and slow orders is generally not monotone with respect to the variance of the demand.

4.5.3 Effect of Leadtime

As mentioned earlier in §4.2, there has not been any discussion in the literature on problems involving random supplies with general leadtimes due to the analytical difficulty. Nevertheless, we attempt to extend our understanding of the leadtime effect by considering the case when the fast supplier has an L-period leadtime and the slow one has an $L+1$-period leadtime. We shall also assume that the capacities are observed one period after placing the corresponding orders. Under this assumption, the results derived in §4.4 continue to hold for general $L > 0$ (see Remark C.0.1 in Appendix).

Table 4.4: The effect of leadtime on order allocation

	Limited Deterministic Capacity			Unlimited Random Capacity			Limited Random Capacity		
	$\bar{q}_f$	$\bar{q}_s$	λ^f	$\bar{q}_f$	$\bar{q}_s$	λ^f	$\bar{q}_f$	$\bar{q}_s$	λ^f
$L=1$	22.43	7.79	74.23 %	12.29	31.22	28.25 %	13.42	30.03	30.88 %
2	25.84	34.93	42.52 %	22.70	64.29	26.10 %	35.00	35.00	50.00 %
3	35.00	35.00	50.00 %	32.79	97.68	25.13 %	35.00	35.00	50.00 %
4	35.00	35.00	50.00 %	42.87	131.06	24.65 %	35.00	35.00	50.00 %
5	35.00	35.00	50.00 %	52.71	164.68	24.25 %	35.00	35.00	50.00 %

Table 4.4 demonstrates that a higher leadtime L leads to a reduced share of the fast order λ^f when the supply is unlimited and random. This is because the one-period leadtime advantage of the fast supplier loses significance when both suppliers have longer leadtimes. A similar effect is also observed under limited deterministic capacities, however only temporarily. Specifically, when the leadtimes increase, the share of the fast order first decreases and then quickly recovers. This increase in λ^f is due to capacity exhaustion. Because increased leadtimes imply higher exposure to demand variability, the desired level of safety stock increases. When the supply capacities are limited, this eventually induces the need to have both suppliers fully utilized. Consequently, the share of the fast supplier λ^f goes up to obtain a 50–50 split

between the two suppliers. This split is more quickly reached when suppliers have limited and random capacities as seen in Table 4.4.

4.6 Concluding Remarks

In this study, we attempt to understand the joint effect of cost, reliability and delivery leadtime on a firm's multiple sourcing strategy. Analyzing these key dimensions of supplier characteristics in a unified framework leads to some interesting observations. First, we show that it may be optimal to order exclusively from a supplier who is inferior in all of the three dimensions. This is driven by capacity limitation or uncertainty associated with the fast supplier. In particular, when the on-hand inventory level is high enough to meet the demand in the current period, a fast order is not needed. However the stockout risk in the next period remains an issue due to the limited or uncertain capacity of the fast supplier. A slow order placed in the current period can effectively mitigate this risk even if it is more expensive and unreliable than the fast. Thus, the firm may source only from an inferior slow supplier when the on-hand inventory is high enough. Second, we observe that an inferior supplier may actually be the primary source of procurement. Due to its unique risk mitigation benefit, when the supply capacity is very restrictive or highly uncertain, firms should order primarily from the slow supplier even if it does not outperform the fast in cost or reliability. These observations highlight the important role the slow supplier plays when sourcing from multiple unreliable suppliers.

To conclude, we discuss the limitations of the model and point out possible avenues for future research. Our model has assumed one-period delay in observing supply capacities. In reality, this observation may be obtained any time from ordering to delivery, i.e., with an m-period delay for $1 \leq m \leq L$. For a general m, the problem becomes more complex due to extended state space. However, we believe that the major insights should carry through, because the advantage of an early slow order in mitigating future stockouts is preserved. An interesting issue to explore is how supply visibility, measured in terms of m, impacts the allocation between the fast and slow suppliers. As in many earlier studies, we have refined our discussion to consecutive leadtimes. A more realistic model should allow for an arbitrary leadtime difference. This again imposes an analytical difficulty and one has to identify efficient heuristic policies. Intuitively, a smaller quantity should be allocated to a slow supplier that has a significantly longer leadtime. If such a notion can be validated, then an important research question is how large the leadtime difference should be to offset the advantage of the slow supplier, leading to a larger allocation to the fast supplier.

Chapter 5

Conclusion

This dissertation explores strategies to mitigate the risks associated with operational and strategic decisions of a firm, particularly focusing on project management, platform development and procurement decisions. In the first essay, we develop methods to evaluate risky capital investment projects that involve managerial flexibility and we illustrate these methods with an alternative energy investment project. In the second essay, we build a strategic model to understand the role of product development decisions as a means to capture demand in two-sided markets. Finally, in the third essay we investigate the value of integrating leadtime flexibility and supply diversification to mitigate supply and demand risks. Our results demonstrate the importance of an integrated approach to risk management.

In the first essay we develop two simulation-based methods to evaluate capital investment projects that involve managerial flexibility. We propose using system dynamics simulations to model the project uncertainty in order to increase the realism of the project model. The methods we develop are based upon first formulating a system dynamics model of the project and then transforming the cash flow data obtained from the model into a decision tree. The

SD-based decision tree approach employs the bracket median approximation technique for this transformation and then evaluates the decision tree using a risk-adjusted discount rate. This is a naive approach to valuation since the same risk-adjusted discount rate is used for the project with and without options regardless of the changing risk character. The diffusion approximation approach overcomes this issue by approximating the cash flow uncertainty generated by the system dynamics model with a binomial decision tree and then evaluating the tree using risk-neutral valuation.

In the literature, it has been argued that there has existed a fundamental trade-off between “detail complexity” and “dynamic complexity”: Financial theory has tended to sacrifice detail complexity, the fidelity of a model at a detailed level, whereas decision analysis has often focused on detail complexity at the expense of keeping some model dynamics unrealistically simple, for example by using a single risk-adjusted discount rate (Smith 1999). In this essay, we show that by using a system dynamics model as an input to evaluating managerial flexibility, it is possible to improve this trade-off between dynamic and detail complexity.

In the second essay, we examine the development of product platforms in markets that exhibit strong cross-network externalities. In many cases, manufacturers of product platforms in these two-sided markets face a trade-off between developing a high performance platform that matches the end-user’s preferences and sacrificing some of those preferences in exchange for improved or less costly third party development capabilities. We use a strategic model to gain intuition about how to make such trade-off decisions under competition and show that conventional wisdom about product development decisions may be misleading in the presence of strong cross-network externalities.

We first characterize content-driven and performance driven markets. The main difference between the two is that end-users highly value content availability in the former and platform performance in the latter. Moreover, in performance-driven markets the platforms are differentiated enough to appeal to different segments of end-users alleviating the intensity of competition. Our results suggest that conventional wisdom may be especially misleading in a content-driven market. For instance, one would expect to see more aggressive investment in product performance if the intensity of competition goes up. However, it turns out in a content-driven market a better strategy for platform providers is to decrease the investment in platform performance and to provide greater content availability instead. More importantly, we show that contrary to the conventional wisdom about "winner-take-all" markets, heavily investing in the core performance of a platform does not always yield a competitive edge when there are strong cross-network effects. In particular, offering a platform with lower performance but greater availability of content may be a better strategy in a content-driven market if platforms are price takers.

The analysis in the second essay focuses on static games of competition between platforms. An interesting direction for future research would be analyzing performance investment strategies of an incumbent and an entrant in a dynamic framework to better understand the extent of installed base advantage. Also, throughout the analysis we assume that content developers can produce content for multiple platforms as long as there is profit to be made, however this does not hold for industries where exclusivity deals bind the content developers to a single platform. If end-users and content developers cannot be affiliated with more than one platform, network effects would have a stronger impact on the market outcome. Our preliminary analysis on this framework shows that the main results for the performance investment strategy continue to hold qualitatively. However, a detailed comparison between the two frameworks is left as future research.

In the third essay, we explore the joint effect of cost, reliability and delivery leadtime on a firm's multiple-sourcing strategy. Specifically, we study dual-sourcing from unreliable suppliers with consecutive delivery leadtimes. We first establish that base-stock policy is generally suboptimal for this procurement problem. Instead, the optimal ordering policy is a threshold policy meaning that there exist reorder points for each supplier such that a positive order is issued if and only if the net inventory level is below the reorder point. We also show that in compensating one unit reduction in net inventory, the total ordering quantity increases by at least one unit, indicating a tendency to mitigate the delivery uncertainty. In other words, diversifying the sourcing profile not only allows for splitting the bet between the two suppliers to share the risk, but also increases the total bet to ensure a certain post-delivery stock level.

Previous results in the dual sourcing and supply diversification literatures suggest that if only one supplier order is issued, then that supplier must be superior in at least one of the dimensions, cost, leadtime or reliability. Interestingly, in our model it can be optimal to order exclusively from a supplier who has a higher cost, longer leadtime and lower reliability than the other. This can be a consequence of two factors. The first factor is the capacity limitation of the fast supplier. When the on-hand inventory is high enough to avoid a fast order in the current period, the firm may be better off ordering some units from a more expensive slow supplier to hedge against the stockout risk in the future periods. The second factor is the unreliability of the fast supply, which can be mitigated by splitting the bet between the fast and the slow suppliers. Specifically, the slow order issued in the current period shares the risk with the fast order in the next period to meet the demand in that period.

We further examine how the firm should allocate the total order between the fast and slow suppliers, who we assume are identical except for their leadtimes. Our analysis suggests that when the fast supplier has ample and reliable capacity, the slow supplier is mainly used as a backup. However, the situation is reversed if the fast supplier's capacity is restrictive or uncertain. In particular, the firm uses the slow supplier as the primary source of procurement even if the slow supplier does not offer any cost or reliability advantage. These observations highlight the importance of incorporating cost, leadtime and reliability in a unified framework to evaluate supplier selection strategies.

Our model has a few simplifying assumptions. First, we assume that there is a one-period delay in observing supply capacities. In reality, capacity uncertainty may resolve any time from ordering to delivery. Allowing for a more general structure of supply visibility would extend the state space and make the problem more complex. However, we believe that the main insights should carry through, as the unique risk sharing benefit of the slow supplier is preserved. Nonetheless, exploring the effects of supply visibility would be an interesting avenue for future research. Second, as in many earlier studies, we assumed a one-period leadtime difference between the two suppliers. An arbitrary leadtime difference would be more realistic, albeit difficult to analyze. One would expect that the allocation to the slow supplier would be reduced if its leadtime is significantly longer. An interesting research question is to explore how large a leadtime difference still preserves the slow supplier's role as a primary source of procurement.

Overall, this dissertation contributes to the existing operations management literature in two ways. First, it illustrates the role of effective risk mitigation through integrating the operational strategies of leadtime flexibility and supply diversification as well as through recognizing managerial flexibility. Second, it highlights the importance of leveraging third-party content development while making technology investment decisions for platforms in two-sided markets. Integrating risk management and operational decisions and the application of this framework to two-sided markets are both promising research directions. We strongly believe that our findings and analyses can be extended to develop new managerial insights for many other issues affecting today's businesses.

This dissertation encompasses a variety of topics in the area of operations management. Exploring these different areas of research has been a motivating experience; however, it also proved to be challenging. The challenge mainly lies in the fact that each area requires mastering a specialized set of tools and techniques. As a result, getting enough exposure to the fundamentals of a domain can be quite demanding. Drawing analogies from operations and risk management, one can recognize the following research diversification trade-off. A very narrow research focus may lead to a difficulty in generating new ideas, which is a risk that can be mitigated by a diversified portfolio of research interests. Yet, a too-diversified researcher cannot effectively take advantage of "economies of scale" in research productivity. One lesson I have learned during the process of writing this dissertation is that for a novice researcher, the latter is likely to be a bigger risk. Thus, for my future studies I am planning to narrow down my research interests. Specifically, my primary focus will be studying two-sided markets from the perspective of operations and risk management.

An operations management perspective can contribute considerably to the study of two-sided markets. The presence of cross-network effects often requires different strategies for two-sided markets than those developed for traditional products and services. The growing literature in this area, however, has been mostly confined to two-sided pricing strategies. Thus, there is more work to be done to improve our understanding of product development, procurement or supply chain design in a two-sided market. The second essay in this dissertation is an attempt in this direction.

In particular, the primary focus of my future research agenda is risk mitigation in two-sided markets. Mainly because it is an emergent field, the study of two-sided markets has been limited to a deterministic framework for the most part. However, an operations management perspective to two-sided markets would be incomplete without recognizing the need to mitigate operational risks. For instance, as highlighted in my third essay, supplier selection and procurement decisions are profoundly affected by the need to mitigate demand and supply risks. It is an open question, though, how such decisions should be made in a two-sided market.

In terms of application areas, a particularly appealing domain is the Smart Grids. The Smart Grid of the future is likely to provide consumers with access to many applications designed to better manage electricity consumption. These applications may have implications similar to the role "apps" play for smartphones. Thus, the two-sided market perspective may be useful in understanding how Smart Grids should be managed to get both the application developers and the consumers on board. Naturally, this is only one facet of the problem. Managing a Smart Grid involves an exceptionally difficult task of matching supply and demand, for which the best practices of operations and risk management are needed. In addition to my training in operations management, the background I acquired while building the alternative energy technologies investment model for the first essay will provide a foundation for my interests in this domain.

The share of two-sided markets in the global economy is rapidly growing. Some of the most important industries are based on this business model, including video-game platforms, credit cards, cell phones, auction sites and operating systems. This provides further basis for research efforts to improve our understanding of two-sided markets; an area where my core research interests lie.

Appendices

Appendix A

Evaluating System Dynamics Models of Risky Projects Using Decision Trees and Real Options Theory

A.1 Fossil Fuel (Natural Gas) Model

The fossil fuel model is developed to obtain a plausible causal model of the fossil fuel price, which is the major uncertainty the alternative energy technologies (AETs) face. We used natural gas data to calibrate the model, since it is mainly the natural gas price that determines the electricity price and thus the profitability of AETs. Therefore, we will describe the model as a natural gas model even though one can safely read it as a generic model for fossil fuels.

A.1.1 Reserves Sector

The total quantity of fossil fuel is divided into three stocks: undiscovered resources, identified reserves and cumulative production. Exploration activities result in discovery of fossil fuel, which depletes *Undiscovered Resources* and fills in the *Identified Reserves.* As natural gas is produced, *Identified* Reserves are depleted. Yet, only a fraction of reserves in the *Identified Reserves* is technically recoverable ($TechRC$).

$$TechRC = Identified\ Reserves * fraction\ recoverable \tag{A.1}$$

Fraction recoverable can be increased with investment in production technology as will be discussed in Section A.1.7.

Potential production rate is constrained by the technically recoverable reserves and production capacity. Moreover, it is assumed that the industry has a desired level of "reserves to production ratio" (R/P ratio), which represents how many years' worth of fossil fuel will be kept in "inventory". One such reasonable level is keeping approximately 10 years' worth of reserves in inventory. Hence, production rate is constrained not only with technically recoverable reserves and production capacity but also the desired level of R/P ratio. Finally, production yield is stochastic: It might be lower than the intended level but cannot be higher, as the upper limit is constrained by technically recoverable reserves and the desired R/P ratio. This stochasticity is modeled via a multiplicative effect formulation *Production Disruptions*. *Production Disruptions* is a combination of *ordinary disruptions*, (minor random jumps in the supply) and *supply shock*, a major disruption that tracks the major oil crisis back in 1970s.

$$\begin{aligned} &Potential\ Production\ Rate = \\ &min\{\frac{TechRC}{Desired\ R/P\ Ratio} * production\ disruptions, Production\ Capacity\} \end{aligned} \tag{A.2}$$

Production rate is the minimum of the *Potential Production Rate* and the *Production Demand*.

$$ProductionRate = min\{PotentialProduction, ProductionDemand\} \quad \text{(A.3)}$$

A.1.2 Exploration Sector

Just like the production sector, only a fraction of potential resources in the *UndiscoveredResources* is technically discoverable (*TechDS*) with the current technology. *Fraction discoverable* (*frDis*) can be increased with investment in exploration technology.

$$TechnDS = UndiscoveredResources * frDis \quad \text{(A.4)}$$

As existing resources diminish, it becomes more expensive to discover new fields. Hence, the productivity of investment in exploration diminishes. The productivity is assumed not to fall below a certain level; hence, the effect of diminishing resources does not exceed a certain threshold, *MaxEffect*. To avoid division by zero error in the extreme case scenarios, the effect is formulated as:

$$\begin{aligned} &effect\ of\ availability \quad \text{(A.5)} \\ &= min\{IFTHENELSE(RR < 1e-009, MaxEffect, RR^{-0.9}), \\ &MaxEffect\} \end{aligned}$$

where

$$\begin{aligned} &RR = IFTHENELSE(TechDS * InitialDiscovered > 0, \\ &TechDS/InitialDiscovered, 0) \quad \text{(A.6)} \end{aligned}$$

Hence, the unit exploration cost (UEC) is:

$$UEC = Initial\, UEC * effect\; of\; availability \tag{A.7}$$

Exploration activities take time. Each time unit, only a fraction of the investment's potential discovery rate is revealed. To account for this, we model *Investment in Exploration (InvE)* as a stock that fills with new investment in exploration and drains when a discovery is actualized.

$$actualized\, investment = Total\; InvE/discovery\; time \tag{A.8}$$

With the *Unit Exploration Cost (UEC)*, this actualized investment rate results in a discovery rate as follows:

$$Discovery\; Rate = actualized\, investment/UEC \tag{A.9}$$

A.1.3 Demand Sector

Production demand has two determinants: Base demand rate and gas price. *Base demand rate (BaseDR)* grows continuously due to population increase that surpasses any possible decrease in energy intensity.

$$growthBaseD = BaseDR * DemandGrowthRatio \tag{A.10}$$

Yet, increasing price has a negative effect on demand, even though in the short run fossil fuel demand is highly inelastic. The effect of price on demand has two components: *Short Term Effect of Price on Demand (STeff)* and *Effect of Conservation (EffCon)*. The former is based on the ratio of the spot price to the expected value of gas price. The expected value is updated when a trend in gas price becomes permanent.

$$\begin{aligned} &change\ in\ exp.\ gas\ price = \\ &(Gas\ Price - Average\ Gas\ Price)/adjustment\ time \end{aligned} \tag{A.11}$$

Further,

$$GasPriceRatio = Gas\ Price/Expected\ Gas\ Price \tag{A.12}$$

$$STeff = Gas\ price\ ratio^{-elasticity} \tag{A.13}$$

Effect of conservation on demand (EffCon) is observed after a delay.

$$EffCon = DELAY3(GasPriceRatio^{C}, conservation\ time) \tag{A.14}$$

Moreover, if the gas price goes beyond a certain price threshold, it will be substituted by other energy resources, which is modeled by *Effect of Switching to Alternative Resources (EffSwitch)*. It is a delayed effect of *price ratio*, the ratio of Gas Price to the price threshold.

$$EffSwitch = f(DELAY3(price\ ratio, switching\ time)) \tag{A.15}$$

where f is described by Figure A.1. Hence, the *Indicated Production Demand (InProdDem)* is computed as:

$$InProdDem = BaseDR * STeff * EffCon * EffSwitch \tag{A.16}$$

Finally, production demand is adjusted every time unit:

$$\Delta ProductionDemand = \frac{InProdDem - Production\ Demand}{demand\ adjustment\ time} \tag{A.17}$$

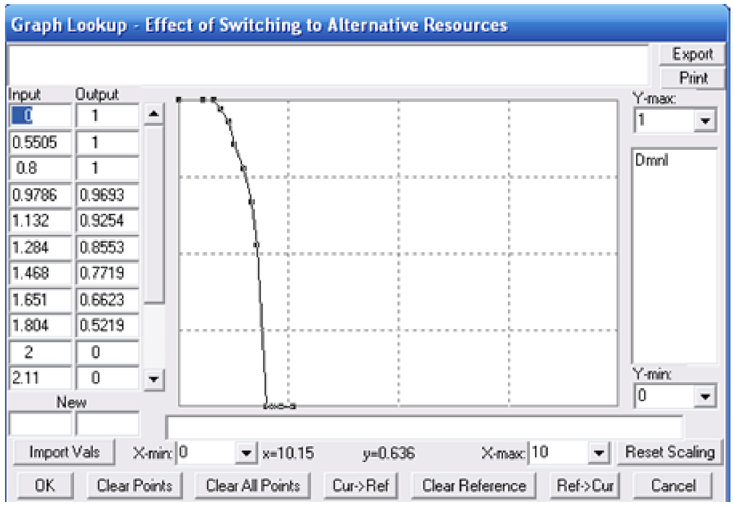


Figure A.1: Graphical Function for the Effect of Switching to Alternative Resources

A.1.4 Gas Price Sector

Three major factors determine the gas price: Demand, supply and total unit production cost, which includes exploration and production costs. The price level that would be obtained when supply and demand is perfectly balanced is called *Unrestricted Gas Price*. This price level is a function of *Unit Production Cost (UPC)* and *Unit Exploration Cost (UEC)*. It is assumed that there is a desired margin the industry would like to achieve. This level determines the *Unrestricted Gas Price* together with the total unit cost.

$$Unrestricted\ Gas\ Pr. = Total\ Unit\ Cost * (1 + desired\ margin) \quad (A.18)$$

When demand for natural gas is higher than its supply, the gas price increases. The effect of demand supply balance on gas price is modeled as follows:

$$Demand\ Supply\ Eff = Demand\ Supply\ Ratio^{c1} \tag{A.19}$$

where

$$DemandSupplyRatio = \frac{Production\ Demand}{max\{Potential\ Production, 0.00001\}} \tag{A.20}$$

Hence, gas price is determined as:

$$Gas\ Price = Unrestricted\ Gas\ Pr. * Demand\ Supply\ Eff \tag{A.21}$$

A.1.5 Production Sector

Production Capacity can be thought of as the refinery capacity. It increases with new *production capacity acquisitions* and decreases with *depreciation. Total Investment in Capacity* increases with new capacity investment (in $) and capacity is acquired after a delay, *capacity building time.* In a way, *Actualized Capacity Investment (ActCapInv)* measures how much capacity would be built now with the *Total Investment in Capacity (TotCapInv)* if the capacity cost were 1.

$$ActCapInv = TotCapInv/capacity\ building\ time \tag{A.22}$$

The actual *Production Capacity Acquisition (ProdCapAcq)* is:

$$ProdCapAcq = ActCapInv/Production\ Capacity\ Cost \tag{A.23}$$

Unit Production Cost (UPC) is a function of remaining resources and remaining reserves.

$$EffRemRes = min\{a1, (\frac{TechDS}{InitialUndiscovered})^{-0.5}\} \tag{A.24}$$

The effect of remaining reserves (*EffRemRez*) is analogous. Hence *Unit Production Cost (UPC)* is modeled as:

$$UPC = Initial\ UPC * EffRemRes * EffRemRez \tag{A.25}$$

Finally, *Production Capacity Cost (ProdCapCost)* is a function of remaining reserves.

$$ProdCapCost = Initial\ ProdCapCost * EffRemRez \tag{A.26}$$

A.1.6 Investment in Exploration and Production Sectors

In equilibrium, the industry is assumed to invest in order to replace the discarded reserves.

$$reserve\ replacement\ inv. = UPC * Production\ Rate \tag{A.27}$$

The return on investment, the supply-demand balance and the R/P ratio affect the actual rate of investment, i.e. whether there is an expansion or contraction in the exploration industry.

$$\begin{aligned} &reserves\ expansion/contraction = \\ &EffDemSupOnInv * EffROI * EffRPratio \end{aligned} \tag{A.28}$$

When there is insufficient demand, there is no motivation to replace the discarded reserves. On the other hand, when demand exceeds supply, investment is boosted.

$$EffDemSupOnInv = IFTHENELSE(R > 1, R^{0.7}, R^2) \tag{A.29}$$

where $R = DemandSupplyRatio$. The firms will investigate the expected return on investment (ROI) to finalize their investment decision. Since this is an aggregate model, we do not get into the details of the ROI computations. For simplicity, the ratio of the forecasted unit price and the total unit cost is used to compute an ROI proxy.

$$\begin{aligned} EffROI &= IFTHENELSE(Price/Cost < 1, (Price/Cost)^{p1}, \\ &\quad (Price/Cost)^{p2}) \end{aligned} \tag{A.30}$$

where

$$Price/Cost = \frac{AverageGasPrice}{UPC + UEC} \tag{A.31}$$

Finally, if the R/P ratio is high, there is less motivation to drill for new reserves. R/P ratio is computed as below:

$$RPratio = \frac{Identified\ Reserves}{Average\ Production} \tag{A.32}$$

$$Relative\ RPratio = \frac{RPratio}{Desired\ RPratio} \tag{A.33}$$

The effect of R/P ratio on investment in exploration is:

$$EffRPratio = f(Relative\ RPratio) \tag{A.34}$$

where f is described by Figure A.2. *Investment in Production* is analogous.

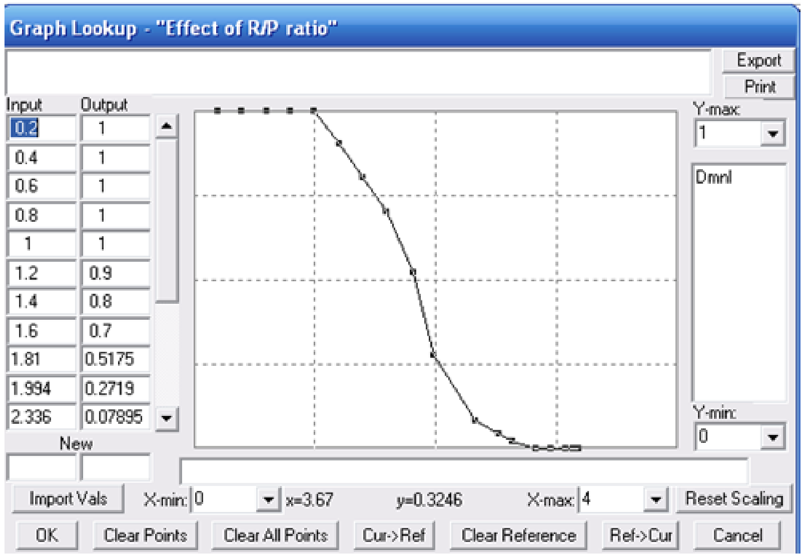


Figure A.2: Effect of R/P Ratio

A.1.7 Technological Advancement

The presence of technological constraints that determine *fraction recoverable* and *fraction discoverable* are already mentioned. These limits can be improved through investment in technology. Improvements are assumed to occur gradually and there is a limit to maximum improvement that can be achieved within a time unit, denoted by *max improvement rate*. Moreover, there is a hypothetical upper limit to technological advancements and the closer the current technology is to this upper limit, the slower the technological improvement rate is. Once we account for these relations, it is the rate of investment in technology that determines the technological improvement rate. This unrestricted rate that can be achieved with a certain investment level is formulated as:

$$IndicatedEffInvOnProdTech =$$
$$max\{1, (InvestmentInProdTech/refInv)^f\} \tag{A.35}$$

Yet, of course, these improvements can only occur after a certain lag, which is modeled as a first-order delay. The actual "unrestricted" effect of investment on technology per unit time is represented by "Effect of Investment on Production Technology" and "Effect of Investment on Exploration Technology"

$$EffInvOnProdTech =$$
$$SMOOTH(IndicatedEffInvOnProdTech, adjTime) \tag{A.36}$$

The change in fraction recoverable (*frRec*) is formulated as:

$$improvement\ in\ prod.\ tech. =$$
$$(1 - frRec) * maxImprovRate * EffInvOnProdTech \tag{A.37}$$

Notice that this formulation guarantees a slower rate of improvement for a given level of investment once we are close to the technological limits. The formulation for fraction discoverable is analogous.

A.2 The Project Model

Firm's capacity acquisition is modeled as a third-order delay. *Alternative Energy Capacity (AEC)* increases as new capacity is acquired and decreases as the existing capacity depreciates. A first-order delay structure is used to represent the rate of depreciation.

$$AECDepreciation = AEC/average\ life\ of\ capacity \tag{A.38}$$

The firm generates electricity at a rate proportional to its installed capacity. Alternative energy technologies like wind and CSP are intermittent; their availability varies in time. The productivity of a wind plant is measured by the capacity factor. Capacity factor (CF) is the ratio of the actual (average) production over the maximum energy the plant would have produced if it had run at full capacity for the whole year. The average capacity factor for wind plants is around 30%. Even though, this average level is known, the actual capacity is stochastic depending on the volatility of wind. Volatility of wind is modeled as a pink noise (autocorrelated noise): White noise is filtered through an exponential smoothing with the correlation time T_s.

$$wind\ volatility = SMOOTH(N(1, \sigma^2), T_s) \tag{A.39}$$

The actual CF is:

$$CF = AverageCF * wind\ volatility \tag{A.40}$$

Hence the average monthly electricity generation rate (*ElecGenRate*) is given by:

$$ElecGenRate = AEC * hoursInMonth * CF \tag{A.41}$$

The supplier enjoys a reduction in its installation costs as the firm acquires more capacity. Hence, the more the generating firm invests, the lower the costs it will face in his future investments. The supplier also takes some advantage of global technological improvements, which is approximated as the impact of a "global" capacity acquisition level (*GlobalCapAcq*). Learning rate is approximated as the sum of the capacity acquisition by the firm and the additional learning through global capacity acquisition adjusted with the weight of global learning:

$$
\begin{aligned}
&LearningRate = CapacityAcquisition + \\
&GlobalCapAcq * weight\ of\ global\ learning
\end{aligned}
\tag{A.42}
$$

Global capacity acquisition (*GlobalCapAcq*) is a function of electricity price and investment costs. As electricity price increases or as the cost of capacity decreases, alternative energy technologies become more competitive. Hence, their global acquisition increases.

$$
\begin{aligned}
&GlobalCapAcq = \\
&refGlobalCapAcq * EffElecPrice * EffInvCosts
\end{aligned}
\tag{A.43}
$$

where

$$EffElecPrice = (\frac{ElectricityPrice}{refElectricityPrice})^{0.5} \tag{A.44}$$

$$EffInvCosts = (InvCostUnitCap/InitialCapCost)^{-0.4} \tag{A.45}$$

The steepness of the learning curve is one of the major uncertainties underlying the alternative energy capacity investment problem. In the literature, learning curve is modeled as:

$$InvCostUnitCap = I * Z^{-\sigma} \tag{A.46}$$

where I is the initial capacity cost, σ is the learning index and Z is the cumulative capacity. Learning index is obtained from the progress ratio, PR, which is empirically determined from the rate of cost reduction whenever cumulative capacity doubles.

$$1 - PR = 2^{-\sigma} \tag{A.47}$$

The formulation is slightly modified in the model. First, the learning curve multiplier (LCM) is based on the cumulative experience relative to the initial level.

$$LCM = \frac{CumExperience}{InitialExperience}^{SteepnessLC} \tag{A.48}$$

Second, it is assumed that investment cost cannot fall below a certain level, *min cost*. Hence, cost of capacity at time t is calculated as:

$$InvCostUnitCap = max\{min\ cost, InitialInvCost * LCM\} \tag{A.49}$$

Cash flow structure works as follows. Revenues accrue from electricity generation.

$$RevenuePerMonth = ElectricityPrice * ElecGenRate \tag{A.50}$$

Electricity price dynamics largely follow gas price dynamics. Hence, we preferred using the following simple representation: Electricity price is at its reference level when gas price is at its reference level. When gas price increases, so does the electricity price (as natural gas is the second most prominent energy source used in electricity generation), but not necessarily linearly.

$$ElectricityPrice = refElectricityPrice * EffGasPrice \tag{A.51}$$

where

$$EffGasPrice = (GasPrice/refGasPrice)^a \tag{A.52}$$

Investment costs incur when new capacity is purchased.

$$InvCost = InvCostUnitCap * CapInvestment \tag{A.53}$$

The firm incurs fixed operating and maintenance cost, which is a function of Alternative Energy Capacity; and variable costs that are a function of electricity generation. The firm also pays the lease rate and taxes. Cost parameters are assumed to be constant and deterministic.

$$\begin{aligned} NetProfitRate &= RevenuePerMonth + PTC \\ &-tax - VariableCosts - FixedOMcosts - LeaseRate \end{aligned} \tag{A.54}$$

Computation of the PTC is slightly involved. If the firm acquires capacity before the expiration date of the PTC, it is eligible for a *credit amount* of 1.5 c/kwh tax credit for a credit duration of 10 years. If the tax the firm has to pay is less than the *Tax Benefit*, the credit is provided for the amount of the tax due. To determine the Tax Benefit at a certain time t, the model checks if the PTC expired for the initial investment and for the expansion investments separately. Because, it might be the case that the credit for the initial investment expired but the credit for the expansion investments have not expired yet; or, the initial investments might occur before the PTC officially expired yet the expansion investments miss that opportunity. If only a part of the electricity generation is eligible for the PTC, the model computes that ratio of revenues that are eligible for the credit. Finally, note that the official expiration date is one of the uncertainties accounted for in the Monte Carlo simulations.

Appendix B

Platform Performance Investment in the Presence of Network Externalities

B.1 Price-Taker Duopoly: Characterization

For the simultaneous-move game described in Section 4.1, the unrestricted platform performance decision at equilibrium is given by

$$\phi^{PT} = \frac{m_p v_\phi - \gamma(t\beta_g - \beta_c(\alpha_g - \gamma))}{\beta_g \gamma v_\phi - 4K\chi}$$

Since the constraint $\phi \geq 0$ is not enforced to obtain this equilibrium, we restrict the parameter space below such that the constraint is trivially satisfied. Note that game developers' margin $\alpha_g - \gamma$ and platform sponsors' adjusted margin $m_p = p - c + \gamma(1 - M)$ are assumed to be nonnegative whereas developers' fixed cost of joining a platform M is assumed to be less than 1, as explained in Section 3 and Section 4.1 respectively. We first divide the parameter space into the following regions:

Region 1: $v_\phi \geq 0$ and $\chi \leq 0$

Region 2: $v_\phi \leq 0$ and $\chi \geq 0$

Region 3: $v_\phi \geq 0$ and $\chi \geq 0$

Region 4: $v_\phi \leq 0$ and $\chi \leq 0$

Next, we exclude Region 3 and Region 4 by showing that under these cases the developer market share N_i, $i = 1, 2$ becomes negative.

Lemma B.1.1. *Region 3 and Region 4 are eliminated.*

Proof. We first show that in Region 3 the developer market share becomes negative at equilibrium and eliminate this region from analysis. Consider a market where $v_\phi = \beta_c - \alpha_c\beta_g \geq 0$ and $\chi = \alpha_c(\alpha_g - \gamma) - t \geq 0$. First note that $\chi = \alpha_c(\alpha_g - \gamma) - t \geq 0$ implies

$$\begin{aligned} t\beta_g - \beta_c(\alpha_g - \gamma) &\leq \alpha_c(\alpha_g - \gamma)\beta_g - \beta_c(\alpha_g - \gamma) = (\alpha_c\beta_g - \beta_c)(\alpha_g - \gamma) \\ &= -v_\phi(\alpha_g - \gamma) \leq 0 \end{aligned}$$

Hence, the numerator of ϕ^{PT} is positive under the assumption $m_p = p - c + \gamma(1 - M) \geq 0$. Accordingly, the condition $\phi^{PT} \geq 0$ requires

$$\beta_g\gamma v_\phi - 4K\chi \geq 0 \tag{B.1}$$

Developer market share at equilibrium is given by

$$N = 1 - M + \frac{\alpha_g - \gamma}{2} - \beta_g\phi^{PT}$$

Hence, the following condition is necessary for developers to enter the market.

$$N = 1 - M + \frac{\alpha_g - \gamma}{2} - \beta_g \frac{m_p v_\phi - \gamma(t\beta_g - \beta_c(\alpha_g - \gamma))}{\beta_g\gamma v_\phi - 4K\chi} \geq 0 \tag{B.2}$$

We want to show that when $v_\phi \geq 0$ and $\chi \geq 0$, (B.2) does not hold. Note that (B.2) implies

$$\beta_g \frac{m_p v_\phi - \gamma(t\beta_g - \beta_c(\alpha_g - \gamma))}{\beta_g \gamma v_\phi - 4K\chi} \leq 1 - M + \frac{\alpha_g - \gamma}{2}$$

Since the denominator of the left hand side is positive by (B.1), we can rewrite the above inequality as

$$\begin{aligned}
&\beta_g v_\phi(p - c + \gamma(1 - M)) - \beta_g \gamma(t\beta_g - \beta_c(\alpha_g - \gamma)) \\
&\qquad \leq (1 - M)\beta_g \gamma v_\phi - (1 - M)4K\chi + \frac{(\alpha_g - \gamma)\beta_g \gamma v_\phi}{2} - 2K\chi(\alpha_g - \gamma)
\end{aligned}$$

Eliminating the term $(1 - M)\beta_g \gamma v_\phi$ from both sides of the inequality yields

$$\begin{aligned}
&(p - c)\beta_g v_\phi - \beta_g \gamma(t\beta_g - \beta_c(\alpha_g - \gamma)) \\
&\qquad \leq -(1 - M)4K\chi + \frac{(\alpha_g - \gamma)\beta_g \gamma v_\phi}{2} - 2K\chi(\alpha_g - \gamma)
\end{aligned}$$

By substituting $v_\phi = \beta_c - \alpha_c \beta_g$ and rearranging the terms, we get

$$\begin{aligned}
&(p - c)\beta_g v_\phi - \frac{(\alpha_g - \gamma)\beta_g \gamma(\beta_c - \alpha_c \beta_g)}{2} - \beta_g^2 \gamma t + \beta_g \gamma \beta_c(\alpha_g - \gamma) \\
&\qquad \leq -(1 - M)4K\chi - 2K\chi(\alpha_g - \gamma) \\
\Leftrightarrow \quad &(p - c)\beta_g v_\phi - \beta_g^2 \gamma t + \beta_g \gamma \beta_c(\alpha_g - \gamma) - \frac{\beta_g \gamma \beta_c(\alpha_g - \gamma)}{2} + \frac{\beta_g^2 \gamma \alpha_c(\alpha_g - \gamma)}{2} \\
&\qquad \leq -(1 - M)4K\chi - 2K\chi(\alpha_g - \gamma) \\
\Leftrightarrow \quad &(p - c)\beta_g v_\phi - \beta_g^2 \gamma t + \frac{\beta_g \gamma \beta_c(\alpha_g - \gamma)}{2} + \frac{\beta_g^2 \gamma \alpha_c(\alpha_g - \gamma)}{2} \\
&\qquad \leq -(1 - M)4K\chi - 2K\chi(\alpha_g - \gamma)
\end{aligned}$$

The right hand side of the last inequality is negative under the aforementioned assumptions $\chi \geq 0$, $M < 1$, $\alpha_g - \gamma \geq 0$. Hence, the left hand side has to be negative. We add and subtract the term $\frac{\beta_g^2\gamma\alpha_c(\alpha_g-\gamma)}{2}$ from the left hand side to get

$$(p-c)\beta_g v_\phi - \beta_g^2\gamma t + \frac{\beta_g\gamma\beta_c(\alpha_g-\gamma)}{2} + \beta_g^2\gamma\alpha_c(\alpha_g-\gamma) - \frac{\beta_g^2\gamma\alpha_c(\alpha_g-\gamma)}{2}$$

Rearranging the terms and substituting $\chi = \alpha_c(\alpha_g-\gamma)-t$ yields

$$\begin{aligned}
&(p-c)\beta_g v_\phi + \beta_g^2\gamma(\alpha_c(\alpha_g-\gamma)-t) + \frac{\beta_g\gamma\beta_c(\alpha_g-\gamma)}{2} - \frac{\beta_g^2\gamma\alpha_c(\alpha_g-\gamma)}{2}\\
&= (p-c)\beta_g v_\phi + \beta_g^2\gamma\chi + \frac{\beta_g\gamma\beta_c(\alpha_g-\gamma)}{2} - \frac{\beta_g^2\gamma\alpha_c(\alpha_g-\gamma)}{2}\\
&= (p-c)\beta_g v_\phi + \beta_g^2\gamma\chi + \frac{\beta_g\gamma(\alpha_g-\gamma)(\beta_c-\beta_g\alpha_c)}{2}\\
&= (p-c)\beta_g v_\phi + \beta_g^2\gamma\chi + \frac{\beta_g\gamma(\alpha_g-\gamma)v_\phi}{2} \qquad \text{(B.3)}
\end{aligned}$$

If the sign of (B.3) is positive, then (B.2) gives a contradiction. The first and the third term of (B.3) are positive since $v_\phi \geq$, whereas the second term is positive since $\chi \geq 0$. This contradicts (B.2). Hence we exclude Region 3 from our analysis. The same result can be shown to hold for Region 4 with a similar proof. □

By excluding Regions 3 and 4, we restrict the parameter space to Regions 1 and 2. In other words, throughout Section 3.4.1.1 we assume that

$$v_\phi\chi = (\beta_c - \alpha_c\beta_g)(\alpha_c(\alpha_g-\gamma)-t) \leq 0 \qquad \text{(B.4)}$$

Next, we restrict Regions 1 and 2 so that $\phi^{PT} \geq 0$ is satisfied.

Region 1: $v_\phi \geq 0$ and $\chi \leq 0$.

In this case, the denominator is always positive. A sufficient condition to make the numerator negative is a high enough m_p. In particular,

$$m_p \geq \frac{\gamma(t\beta_g - \beta_c(\alpha_g - \gamma))}{v_\phi} \tag{B.5}$$

This sufficient condition is hardly restrictive when the competition in the end-user market is intense, i.e. when t is small. Our initial assumption $m_p \geq 0$ usually satisfies the condition in that case. In a market with well-differentiated platforms, however, $m_p \geq 0$ may not be enough to guarantee a nonnegative equilibrium value for the platform performance, especially when β_c is relatively small compared to α_c (i.e $v_\phi \leq 0$.). In other words, we see that in a competitive market a small margin may be enough to motivate the platform sponsor invest in platform performance. However, once the platforms are well-differentiated, a higher margin is required for positive investment in platform performance, especially if end-users' utility from content availability exceeds their utility from platform performance.

Region 2: $v_\phi \leq 0$ and $\chi \geq 0$.

In this case, the denominator is always negative. Just as in Region 1, a sufficient condition to make the numerator negative is a high enough m_p.

$$\begin{aligned} m_p v_\phi - \gamma(t\beta_g - \beta_c(\alpha_g - \gamma)) \leq 0 \\ \Leftrightarrow m_p \geq \frac{\gamma(t\beta_g - \beta_c(\alpha_g - \gamma))}{v_\phi} \end{aligned}$$

The second inequality follows from the fact that $v_\phi \leq 0$.

Throughout the analysis, we assume that the end-user market is covered. To satisfy this assumption, the net utility of the marginal end-user that is indifferent between the two platforms must be nonnegative. Note that in the Hotelling model this marginal end-user is "farthest" away from the platform he is affiliated with. Accordingly, the net utility of the marginal end-user from purchasing platform i, $i = 1, 2$ is given by

$$u_i^m = \alpha_c N_i + \beta_c \phi_i - p - tD_i \tag{B.6}$$

At equilibrium, the end-user market is split equally between the platforms. Substituting (B.2) for N_i, 1/2 for D_i, and ϕ^{PT} for ϕ_i

$$u^m = \alpha_c(1 - M + 1/2(\alpha_g - \gamma) - \beta_g \phi^{PT}) + \beta_c \phi^{PT} - p - t/2$$

Rearranging the terms we get

$$\begin{aligned} u^m &= \alpha_c(1 - M + 1/2(\alpha_g - \gamma)) + (\beta_c - \alpha_c \beta_g)\phi^{PT} - p - t/2 \\ &= \alpha_c(1 - M + 1/2(\alpha_g - \gamma)) + v_\phi \phi^{PT} - p - t/2 \\ &= \alpha_c(1 - M) + v_\phi \phi^{PT} - p + \chi/2 \end{aligned} \tag{B.7}$$

Market coverage assumption requires (B.7) to be nonnegative. In other words we need

$$p - \chi/2 \leq \alpha_c(1 - M) + v_\phi \phi^{PT} \tag{B.8}$$

Note that low M and high α_c makes it likely that the market coverage assumption (B.8) to be satisfied. Further, low M implies a higher m_p, which helps to satisfy (B.5) as well.

Finally, sufficient conditions for optimality require the following derivative to be nonpositive.

$$\frac{\partial \Pi^2}{\partial^2 \phi} = -2K + \frac{v_\phi \gamma(\beta_g(\chi - t) + \beta_c(\alpha_g - \gamma))}{2\chi^2} \tag{B.9}$$

Note that the denominator of the second term is always positive. Accordingly, to satisfy the sufficient condition for optimality, K (platform development cost per unit performance squared) must be sufficiently high. In particular,

$$K \geq \frac{v_\phi \gamma(\beta_g(\chi - t) + \beta_c(\alpha_g - \gamma))}{4\chi^2} \tag{B.10}$$

To sum up, throughout Section 4.1 we assume that the parameters satisfy (B.4), (B.5),(B.8) and (B.10) together with the initial assumptions $\alpha_g - \gamma \geq 0$, $M < 1$ and $m_p \geq 0$.

B.2 Price-Setting Duopoly: Characterization

We see that at equilibrium the platforms play one of the two pooling strategies described in Lemma 3.4.2 depending on the market parameters. In particular, either both platforms make zero investment in platform performance or they make a positive investment and adjust the price accordingly. We first characterize the parameter region where the high performance equilibrium holds. Naturally, the first condition is

$$\beta_c \geq \beta_g(\alpha_c + \gamma) \tag{B.11}$$

The following conditions must be met to satisfy the sufficient conditions for optimality:

$$\frac{\partial^2 \Pi}{\partial p^2} = \frac{-2t + (\alpha_g - \gamma)(2\alpha_c + \gamma)}{2(t - \alpha_c(\alpha_g - \gamma))^2} \leq 0 \tag{B.12}$$

$$\frac{\partial^2 \Pi}{\partial \phi^2} = -2K + \frac{\gamma v_\phi(-2t\beta_g + (\beta_c + \alpha_c\beta_g)(\alpha_g - \gamma))}{2(t - \alpha_c(\alpha_g - \gamma))^2} \leq 0 \tag{B.13}$$

$$\frac{\partial^2 \Pi}{\partial p^2}\frac{\partial^2 \Pi}{\partial \phi^2} - (\frac{\partial^2 \Pi}{\partial p \partial \phi})^2 = \frac{4K(2t - (\alpha_g - \gamma)(2\alpha_c + \gamma)) - v_{PS}^2}{4(t - \alpha_c(\alpha_g - \gamma))^2} \geq 0 \tag{B.14}$$

From (B.14), the following necessary condition arises

$$2t - (\alpha_g - \gamma)(2\alpha_c + \gamma) \geq \frac{v_{PS}^2}{4K} \tag{B.15}$$

Note that when (B.15) is satisfied, so is (B.12) because (B.15) implies that the numerator of (B.12) is negative. Further, a significantly high platform differentiation guarantees to satisfy (B.13). To see this, note that $v_\phi = \beta_c - \beta_g\alpha_c \geq 0$ due to the fact that v_{PS} is positive. Hence, if t is high enough to make the second term $(-2t\beta_g + (\beta_c + \alpha_c\beta_g)(\alpha_g - \gamma))$ negative, (B.13) is satisfied. Note that this is only a sufficient condition. Even if t is not high enough, (B.13) would still be satisfied with a sufficiently high K.

Finally, throughout the analysis we assume that end-user market is covered, which requires

$$\begin{aligned} u^m &= \frac{v_{PS}^2}{4K} - c - 3/2\, t \\ &+ \alpha_c(1 - M + 3/2\, (\alpha_g - \gamma)) + \gamma(1 - M + \alpha_g - \gamma) \geq 0 \end{aligned} \tag{B.16}$$

This condition puts an upperbound on t and c. In particular,

$$3/2\ t + c \leq \frac{v_{PS}^2}{4K} + \alpha_c(1 - M + 3/2\ (\alpha_g - \gamma)) + \gamma(1 - M + \alpha_g - \gamma) \quad \text{(B.17)}$$

Accordingly, throughout Section 4.2 we assume that (B.13), (B.15) and (B.17) are satisfied, together with our initial assumptions $M < 1$ and $\alpha_g - \gamma \geq 0$.

B.3 Proofs

Proof of Proposition 3.4.1.

i) It suffices to check the sign of the following derivative:

$$\frac{\partial \phi^{PT}}{\partial \beta_c} = \frac{-\chi(\beta_g^2\gamma^2 + 4K(m_p + \gamma(\alpha_g - \gamma)))}{[-4K\chi + \beta_g\gamma v_\phi]^2} \quad \text{(B.18)}$$

It is easy to see that (B.18) is nonpositive if $\chi \geq 0$ under the assumption $m_p \geq 0$. By Lemma B.1.1, $\chi \geq 0$ also requires $v_\phi \leq 0$ concluding the proof.

ii) Note that the degree of competition in a market increases when t, the product differentiation between the platforms, decreases. Hence, it suffices to confirm that ϕ^{PT} may increase with t by checking the sign of the following derivative:

$$\frac{\partial \phi^{PT}}{\partial t} = \frac{-v_\phi(\beta_g^2\gamma^2 + 4K(m_p + \gamma(\alpha_g - \gamma)))}{[-4K\chi + \beta_g\gamma v_\phi]^2} \quad \text{(B.19)}$$

The denominator of (B.19) is trivially positive. The numerator is positive when $v_\phi \leq 0$ under the assumption $m_p \geq 0$. By Lemma B.1.1, $v_\phi \leq 0$

also requires $\chi \geq 0$, concluding the proof. Hence, when v_ϕ is negative and χ is positive, platform performance decreases with increasing degree of competition.

iii) It suffices to check the sign of the following derivative:

$$\frac{\partial \phi^{PT}}{\partial \alpha_g} = \frac{v_\phi(\beta_g \beta_c \gamma^2 + 4K\alpha_c m_p + 4Kt\gamma)}{[-4K\chi + \beta_g \gamma v_\phi]^2} \tag{B.20}$$

The denominator of (B.20) is trivially positive. The numerator is negative if $v_\phi \leq 0$ under the assumption $m_p \geq 0$, in which case the optimal performance decreases with the game price α_g. By Lemma B.1.1, $v_\phi \leq 0$ also requires $\chi \geq 0$, concluding the proof.

iv) It suffices to check the sign of the following derivative:

$$\frac{\partial \phi^{PT}}{\partial \alpha_c} = -\frac{(t\beta_g - \beta_c(\alpha_g - \gamma))(\beta_g^2 \gamma^2 + 4K(m_p + \gamma(\alpha_g - \gamma)))}{[-4K\chi + \beta_g \gamma v_\phi]^2} \tag{B.21}$$

It is easy to see that (B.21) is nonnegative when $t\beta_g - \beta_c(\alpha_g - \gamma) \leq 0$ under the assumption $m_p \geq 0$.

□

Proof of Proposition 3.4.3. First note that when $v_{PS} = \beta_c - \beta_g(\alpha_c + \gamma) \leq 0$, both platforms choose the minimum platform performance which is assumed to be zero. In that case, equilibrium performance is insensitive to changes in market parameters as long as v_{PS} stays negative. On the other hand, when $v_{PS} \geq 0$, platforms play the high performance equilibrium and set $\phi_H^{PS} = v_{PS}/4K$.

i) If $v_{PS} \geq 0$, it suffices to check the sign of the following derivative which is trivially nonnegative.

$$\frac{\partial \phi_H^{PS}}{\partial \beta_c} = 1/4K$$

On the other hand, if $v_{PS} \leq 0$, equilibrium performance does not change with β_c unless the increase in β_c makes v_{PS} positive in which case equilibrium performance increases. Hence, equilibrium performance is non-decreasing in β_c.

ii) An increase in α_c further decreases v_{PS}. Hence, if $v_{PS} \leq 0$, platform performance does not change when α_c increases. On the other hand, if $v_{PS} \geq 0$, it suffices to check the sign of the following derivative which is trivially nonpositive.

$$\frac{\partial \phi_H^{PS}}{\partial \alpha_c} = -\beta_g/4K$$

Hence, equilibrium performance is nonincreasing in α_c.

iii) An increase in γ further decreases v_{PS}. Hence, if $v_{PS} \leq 0$, platform performance does not change when γ increases. On the other hand, if $v_{PS} \geq 0$, it suffices to check the sign of the following derivative which is trivially nonpositive.

$$\frac{\partial \phi_H^{PS}}{\partial \gamma} = -\beta_g/4K$$

iv) An increase in β_g further decreases v_{PS}. Hence, if $v_{PS} \leq 0$, platform performance does not change when β_g increases. On the other hand, if

$v_{PS} \geq 0$, it suffices to check the sign of the following derivative which is trivially nonpositive.

$$\frac{\partial \phi_H^{PS}}{\partial \beta_g} = -(\alpha_c + \gamma)/4K$$

v) It is easy to see that $\partial\phi_H^{PS}/\partial t = \partial\phi_L^{PS}/\partial t = 0$

□

Proof of Proposition 3.4.4. Consider two symmetric platforms simultaneously entering the market. By substituting the equilibrium values for ϕ^{PS}, p^{PS}, N^{PS} and D^{PS} into the profit function (3.6), profit earned by each platform in a price setting duopoly is given by

$$\pi^{PS} = \begin{cases} \frac{4K(2t-(\alpha_g-\gamma)(2\alpha_c+\gamma))-(\beta_c-\beta_g(\alpha_c+\gamma))^2}{16K} & v_{PS} \geq 0 \\ \frac{2t-(\alpha_g-\gamma)(2\alpha_c+\gamma)}{4} & v_{PS} \leq 0, \end{cases} \tag{B.22}$$

Similarly, if $v_\phi\chi \leq 0$, profit earned by each platform in a price-taker duopoly is given by

$$\begin{aligned} \pi^{PT} = 1/2\ (p - c - \frac{2K(m_p v_\phi - \gamma(t\beta_g - \beta_c(\alpha_g - \gamma)))^2}{(\beta_g\gamma v_\phi - 4K\chi)^2}) + \\ 1/4\gamma(2(1-M) + \alpha_g - \gamma - 2\beta_g\frac{m_p v_\phi - \gamma(t\beta_g - \beta_c(\alpha_g - \gamma))}{\beta_g\gamma v_\phi - 4K\chi}) \end{aligned} \tag{B.23}$$

whereas if $v_\phi\chi > 0$,

$$\pi^{PT} = \frac{2(p-c) + \gamma(2(1-M) + \alpha_g - \gamma)}{4} \tag{B.24}$$

where $v_\phi = \beta_c - \alpha_c\beta_g$, $\chi = \alpha_c(\alpha_g - \gamma) - t$ and $m_p = p - c + \gamma(1-M)$.

i) It is easy to show that

$$\frac{\partial \pi^{PS}}{\partial \beta_c} = \begin{cases} \frac{-\beta_c+\beta_g(\alpha_c+\gamma)}{8K} & v_{PS} \geq 0 \\ 0 & v_{PS} < 0 \end{cases} \tag{B.25}$$

whereas

$$\frac{\partial \pi^{PT}}{\partial \beta_c} = \begin{cases} \frac{v_\phi \chi(\beta_g^2\gamma^2+4K(m_p+\alpha_g-\gamma))^2}{2(\beta_g\gamma v_\phi-4K\chi)^3} & v_\phi\chi \leq 0 \\ 0 & v_\phi\chi > 0 \end{cases} \tag{B.26}$$

First, we look at the price-setting duopoly. If $v_{PS} < 0$, profit is insensitive to changes in β_c. When $v_{PS} \geq 0$, the numerator of (B.25) is negative. Since the denominator is trivially positive, this implies $\partial\pi^{PS}/\partial\beta_c$ is negative. In other words, in a price setting duopoly if the end-users' utility from platform performance increases the profit of each platform decreases.

Next, we look at the price-taker duopoly. If $v_\phi\chi > 0$, $\phi^{PT} = 0$ and from $B.26$ profit is insensitive to changes in β_c. On the other hand, when $v_\phi\chi \leq 0$, the numerator of (B.26) is always negative. If $v_\phi \leq 0$, then $\chi \geq 0$; in which case the denominator also becomes negative, making (B.26) positive. Hence, in a content-driven market characterized by $v_\phi \leq 0$ and $\chi \geq 0$, the profit of a price-taker platform may actually increase when the end-users' utility from platform performance increases.

ii) It is easy to show that

$$\frac{\partial \pi^{PS}}{\partial t} = 1/2 \tag{B.27}$$

$$\frac{\partial \pi^{PT}}{\partial t} = \begin{cases} \frac{v_\phi^2(\beta_g^2\gamma^2+4K(m_p+\alpha_g-\gamma))^2}{2(\beta_g\gamma v_\phi-4K\chi)^3} & v_\phi\chi \leq 0 \\ 0 & v_\phi\chi > 0 \end{cases} \tag{B.28}$$

Trivially from (B.27), the profit of a price-setting platform always increases when t increases, in other words, when the degree of competition decreases. To see that π^{PT} may actually increase when the degree of competition decreases, i.e. that (B.28) may be negative, first note that the numerator is trivially positive. By (B.4), we know that $v_\phi \chi \leq 0$. This implies that if $v_\phi \leq 0$, then $\chi \geq 0$; in which case the denominator becomes negative. Hence in a content-driven market characterized by $v_\phi \leq 0$ and $\chi \geq 0$, the profit a price-taker platform may actually increase when the degree of competition increases, concluding the proof.

□

Proof of Proposition 3.4.5.

i) In Proposition 3.4.3, we have shown that ϕ^{PS} decreases with β_g. It suffices to show that p^{PS} can increase with β_g.

$$\frac{\partial p^{PS}}{\partial \beta_g} = \begin{cases} \frac{\gamma(\beta_c - 2\beta_g(\alpha_c+\gamma))}{4K} & v_{PS} \geq 0 \\ 0 & v_{PS} < 0 \end{cases} \tag{B.29}$$

It is easy to see that for $v_{PS} \geq 0$ the numerator of (B.29) is positive when $\beta_c - 2\beta_g(\alpha_c + \gamma) > 0$, in which case p^{PS} increases with β_g even though ϕ^{PS} decreases.

ii) In Proposition 3.4.3, we have shown that ϕ^{PS} decreases with γ when $v_{PS} \geq 0$. It suffices to show that p^{PS} can increase with γ under the same conditions.

$$\frac{\partial \phi^{PS}}{\partial \gamma} = \begin{cases} \frac{\beta_g(\beta_c - \beta_g(\alpha_c + 2\gamma)) - 4K(1 - M - \alpha_c + \alpha_g - 2\gamma)}{4K} & v_{PS} \geq 0 \\ \alpha_c + 2\gamma - 1 + M - \alpha_g) & v_{PS} < 0 \end{cases} \tag{B.30}$$

When $\beta_g(\beta_c - \beta_g(\alpha_c + 2\gamma)) \geq 4K(1 - M - \alpha_c + \alpha_g - 2\gamma)$, the numerator of (B.30) is positive in which case p^{PS} increases with γ even though ϕ^{PS} decreases.

□

Proof of Proposition 3.5.1. Let Platform 1 be the leader and Platform 2 the follower. An analysis of the first order conditions yield the equilibrium values of platform performances as follows:

$$\phi_1^{PSseq} = \frac{-v_{PS}(v_{PS}^2 + K(2(\alpha_g - \gamma)(3\alpha_c + 2\gamma) - 6t))}{2K(-3v_{PS}^2 + 4K(4t - (\alpha_g - \gamma)(4\alpha_c + 3\gamma)))} \tag{B.31}$$

$$\phi_2^{PSseq} = \frac{-v_{PS}(v_{PS}^2 + K((\alpha_g - \gamma)(5\alpha_c + 4\gamma) - 5t))}{K(-3v_{PS}^2 + 4K(4t - (\alpha_g - \gamma)(4\alpha_c + 3\gamma)))} \tag{B.32}$$

i) It is easy to show that

$$\frac{\partial \phi_1^{PSseq}}{\partial \alpha_g} = \frac{v_{PS}(4Kt\gamma + \alpha_c v_{PS}^2)}{(3v_{PS}^2 - 4K(4t - (\alpha_g - \gamma)(4\alpha_c + 3\gamma)))^2}$$
$$= -\frac{\partial \phi_2^{PSseq}}{\partial \alpha_g} = \frac{-v_{PS}(4Kt\gamma + \alpha_c v_{PS}^2)}{(3v_{PS}^2 - 4K(4t - (\alpha_g - \gamma)(4\alpha_c + 3\gamma)))^2}$$

ii) It is easy to show that

$$\frac{\partial \phi_1^{PSseq}}{\partial t} = -v_{PS}\frac{v_\phi^2 + 4K\gamma\alpha_g - 2\beta_g\gamma v_\phi + \gamma^2(\beta_g^2 - 4K)}{(3v_{PS}^2 - 4K(4t - (\alpha_g - \gamma)(4\alpha_c + 3\gamma)))^2}$$
$$= -\frac{\partial \phi_2^{PSseq}}{\partial t} = v_{PS}\frac{v_\phi^2 + 4K\gamma\alpha_g - 2\beta_g\gamma v_\phi + \gamma^2(\beta_g^2 - 4K)}{(3v_{PS}^2 - 4K(4t - (\alpha_g - \gamma)(4\alpha_c + 3\gamma)))^2}$$

concluding the proof. $\square$

Appendix C

Dual Sourcing under Random Supply Capacities: The Role of the Inferior Supplier

Theorem C.0.1. *For each t, the following results hold.*

i) *The objective function $J_t(I; q_f, q_s)$ is continuous in (I, q_f, q_s) and the optimal cost function $V_t(I)$ is continuous in I.*

ii) *The optimal cost function $V_t(I) < \infty$ and $\lim_{I \to \pm\infty} V_t(I) = \infty$*

iii) *For each I, there exist finite upper bounds on the optimal q_f and q_s that minimize $J_t(I; q_f, q_s)$.*

Proof. Part i) follows from a simple inductive argument together with the fact that the one-period cost function is continuous in (I, q_f, q_s).

To see part ii), note that

$$\begin{aligned} V_t(I) &\leq \sum_{s=t}^{T} \alpha^{s-t} [\mathbb{E} H(I - \sum_{s=t}^{T} D_t)] \\ &\leq \sum_{s=t}^{T} \alpha^{s-t} \left[\mathbb{E} c^H \left| I - \sum_{s=t}^{T} D_t \right| \right] < \infty \end{aligned}$$

The first inequality follows from the fact that the optimal cost in period t is less than the cost incurred by ordering nothing over the periods t to T. The second inequality follows from the assumption that $|H_t(x_1) - H_t(x_2)| \leq c^H |x_1 - x_2|$. The last inequality follows from the assumption that $\mathbb{E}D_t$ is finite for each t. The result that $\lim_{I \to \pm\infty} V_t(I) = \infty$ follows directly from the assumption that $\lim_{I \to \pm\infty} H(I) = \infty$, proving Part ii).

Part iii) follows directly from part ii). □

Lemma C.0.2. *Let $\phi(q)$ be a convex function minimized at $q = \tilde{q}$ and let k, k^a and k^b be identically distributed nonnegative random variables. For any given $\theta \in [0,1]$ with $\bar{\theta} = 1 - \theta$, we have the following relations.*

i) *If k^a and k^b are independent, then*

$$\mathbb{E}\phi(\tilde{q} \wedge k) \geq \min_{q^a \geq 0, q^b \geq 0} \{\mathbb{E}\phi(\theta(q^a \wedge k^a) + \bar{\theta}(q^b \wedge k^b))\}.$$

ii) *If $k^a = k^b = k$ almost surely, then for any choice of (q^a, q^b),*

$$\mathbb{E}\phi(\tilde{q} \wedge k) \leq \mathbb{E}\phi(\theta(q^a \wedge k) + \bar{\theta}(q^b \wedge k)).$$

Proof of LemmaC.0.2. We first prove part i). When k^a and k^b are independent copies of k, we have

$$\mathbb{E}\phi(\theta(q \wedge k^a) + \bar{\theta}(q \wedge k^b)) \leq \mathbb{E}[\theta\phi(q \wedge k^a) + \bar{\theta}\phi(q \wedge k^b))] = \mathbb{E}\phi(q \wedge k).$$

The first inequality follows from the convexity of ϕ, and the second follows from the fact that k^a, k^b, and k have the same distribution.

To see part ii), consider possible realizations of $k = k^a = k^b$,

Case 1: If $k < \tilde{q}$, then

$$\theta(q^a \wedge k) + \bar{\theta}(q^b \wedge k) \leq k < \tilde{q}.$$

We must have

$$\phi(\theta(q^a \wedge k) + \bar{\theta}(q^b \wedge k)) \geq \phi(k) = \phi(\tilde{q} \wedge k).$$

Case 2: If $k \geq \tilde{q}$, then by the optimality of $\tilde{q}$, we have

$$\phi(\theta(q^a \wedge k) + \bar{\theta}(q^b \wedge k)) \geq \phi(\tilde{q}) = \phi(\tilde{q} \wedge k).$$

Combining Cases 1 and 2, we conclude that

$$\mathbb{E}\phi(\theta(q^a \wedge k) + \bar{\theta}(q^b \wedge k)) \geq \mathbb{E}\phi(\tilde{q} \wedge k).$$

□

Proof of Proposition 4.4.1. Define the cost function for realized delivery quantities y_f and y_s as follows

$$\pi(I; y_f, y_s) = c_f y_f + c_s y_s + \mathbb{E}H(I + y_f - D) + \alpha \mathbb{E}V_{t+}(I + y_f + y_s - D).$$

Clearly $\pi(I; y_f, y_s)$ is jointly convex in (I, y_f, y_s). Also note that

$$J_t(I; q_f, q_s) = \mathbb{E}\pi(I; q_f \wedge k_f, q_s \wedge k_s).$$

Choose $I_1 < I_2$ and $\theta \in [0,1]$ with $\bar{\theta} = 1 - \theta$. Also define $I = \theta I_1 + \bar{\theta} I_2$. We have

$$
\begin{aligned}
& \theta V_t(I_1) + \bar{\theta} V_t(I_2) && \text{(C.1)} \\
= \quad & \theta J_t(I_1; q_f^*(I_1), q_s^*(I_1)) + \bar{\theta} J_t(I_2; q_f^*(I_2), q_s^*(I_2)) \\
= \quad & \theta \mathbb{E}\pi(I_1; q_f^*(I_1) \wedge k_f, q_s^*(I_1) \wedge k_s) + \bar{\theta} \mathbb{E}\pi(I_2; q_f^*(I_2) \wedge k_f, q_s^*(I_2) \wedge k_s) \\
\geq \quad & \mathbb{E}\pi(I; \theta(q_f^*(I_1) \wedge k_f) + \bar{\theta}(q_f^*(I_2) \wedge k_f), \theta(q_s^*(I_1) \wedge k_s) + \bar{\theta}(q_s^*(I_2) \wedge k_s)).
\end{aligned}
$$

Let $y_f = \tilde{q}_f$ be the minimizer of

$$\phi^f(y_f) = \mathbb{E}\pi(I; y_f, \theta(q_s^*(I_1) \wedge k_s) + \bar{\theta}(q_s^*(I_2) \wedge k_s)).$$

Note that the right-hand side of (C.1) is simply

$$\mathbb{E}\phi^f(\theta(q_f^*(I_1) \wedge k_f) + \bar{\theta}(q_f^*(I_2) \wedge k_f)).$$

By Lemma C.0.2, we conclude that

$$\mathbb{E}\phi^f(\theta(q_f^*(I_1) \wedge k_f) + \bar{\theta}(q_f^*(I_2) \wedge k_f)) \geq \mathbb{E}\phi^f(\tilde{q}_f \wedge k_f).$$

Now define

$$\phi^s(y_s) = \mathbb{E}\pi(I, \tilde{q}_f \wedge k_f, y_s).$$

Let $\tilde{q}_s$ be the minimizer of $\phi^s(y_s)$. Note that

$$\mathbb{E}\phi^f(\tilde{q}_f \wedge k_f) = \mathbb{E}\phi^s(\theta(q_s^*(I_1) \wedge k_s) + \bar{\theta}(q_s^*(I_2) \wedge k_s)).$$

A similar argument as the above shows that

$$\mathbb{E}\phi^s(\theta(q_s^*(I_1) \wedge k_s) + \bar{\theta}(q_s^*(I_2) \wedge k_s)) \geq \mathbb{E}\phi^s(\tilde{q}_s \wedge k_s).$$

We deduce that

$$
\begin{aligned}
\theta V_t(I_1) + \bar{\theta} V_t(I_2) &\geq \mathbb{E}\phi^f(\tilde{q}_f \wedge k_f) \\
&\geq \mathbb{E}\phi^s(\tilde{q}_s \wedge k_s) \\
&= \mathbb{E}\pi(I, \tilde{q}_f \wedge k_f, \tilde{q}_s, \wedge k_s) \\
&= J_t(I; \tilde{q}_f, \tilde{q}_s) \\
&\geq V_t(I).
\end{aligned}
$$

Hence, we conclude the proof. □

PROOF OF PROPOSITION 4.4.2. The proof is based on the convexity of $V_{t+1}(\cdot)$ and $H(\cdot)$ and the first-order conditions for $q_{f,t}$ and $q_{s,t}$ given as follows:

$$
\frac{\partial J_t}{\partial q_{f,t}} = \phi_{f,t}(I, q_{f,t}, q_{s,t}) \bar{G}_f(q_{f,t}) = 0, \tag{C.2}
$$

$$
\frac{\partial J_t}{\partial q_{s,t}} = \phi_{s,t}(I, q_{f,t}, q_{s,t}) \bar{G}_s(q_{s,t}) = 0, \tag{C.3}
$$

where $G_{j,t}(\cdot)$ is the distribution of $k_{j,t}$ with $\bar{G}_{j,t}(x) = 1 - G_{j,t}(x)$, $j = f, s$ and

$$
\begin{aligned}
\phi_{f,t}(I, q_{f,t}, q_{s,t}) &= c_f + \mathbb{E}H'(I + q_{f,t} - D) + \alpha \mathbb{E}V'_{t+1}(I + q_{f,t} + q_{s,t} \wedge k_s - D), \\
\phi_{s,t}(I, q_{f,t}, q_{s,t}) &= c_s + \alpha \mathbb{E}V'_{t+1}(I + q_{f,t} \wedge k_f + q_{s,t} - D).
\end{aligned}
$$

We first show that $q^*_{f,t}(I) \leq q^*_{f,t}(I - \delta)$ for $\delta > 0$. Suppose, on the contrary,

$q^*_{f,t}(I) > q^*_{f,t}(I-\delta)$. The first-order condition with respect to $q_{f,t}$ requires:

$$\begin{aligned}
0 &\geq \phi_f(I; q^*_{f,t}(I), q^*_{s,t}(I)) - \phi_f(I-\delta; q^*_{f,t}(I-\delta), q^*_{s,t}(I-\delta)) \qquad \text{(C.4)}\\
&= \mathbb{E}H'(I + q^*_{f,t}(I) - D) - \mathbb{E}H'(I - \delta + q^*_{f,t}(I-\delta) - D)\\
&+\alpha \max\{\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) - D), A(I)\}\\
&-\alpha \max\{\mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) - D), A(I-\delta)\}
\end{aligned}$$

where

$$A(I) = \mathbb{E}\min\{\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + q^*_{s,t}(I) - D), \mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + k_s - D)\}$$

Note that by the convexity of $H(\cdot)$ and $V_{t+1}(\cdot)$, the assumption $q^*_{f,t}(I) > q^*_{f,t}(I-\delta)$ implies:

$$\begin{aligned}
&\mathbb{E}H'(I + q^*_{f,t}(I) - D) - \mathbb{E}H'(I - \delta + q^*_{f,t}(I-\delta) - D) \geq 0\\
&\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) - D) \geq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) - D)\\
&\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + k_s - D) \geq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) + k_s - D)
\end{aligned}$$

Hence, to satisfy (C.4), the following must hold:

$$\begin{aligned}
&\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + q^*_{s,t}(I) - D)\\
&\qquad \leq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - D) \qquad \text{(C.5)}
\end{aligned}$$

By the convexity of $V_{t+1}(\cdot)$, this is equivalent to

$$q^*_{f,t}(I) + q^*_{s,t}(I) \leq q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - \delta \qquad \text{(C.6)}$$

Note that if $q^*_{s,t}(I) \geq q^*_{s,t}(I-\delta)$, C.6 would be a contradiction under the assumption $q^*_{f,t}(I) > q^*_{f,t}(I-\delta)$. Now consider the first-order condition with respect to the slow order $q_{s,t}(I)$.

$$\begin{aligned}
&0 \leq \phi_s(I, q^*_{f,t}(I), q^*_{s,t}(I)) - \phi_s(I-\delta, q^*_{f,t}(I-\delta), q^*_{s,t}(I-\delta)) \qquad \text{(C.7)}\\
&= \alpha \max\{\mathbb{E}V'_{t+1}(I + q^*_{s,t}(I) - D), B(I)\}\\
&-\alpha \max\{\mathbb{E}V'_{t+1}(I-\delta + q^*_{s,t}(I-\delta) - D), B(I-\delta)\}
\end{aligned}$$

where

$$B(I) = \mathbb{E}\min\{\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + q^*_{s,t}(I) - D), \mathbb{E}V'_{t+1}(I + k_f + q^*_{s,t}(I) - D)\}$$

But given (C.5), (C.7) requires

$$q^*_{s,t}(I) \geq q^*_{s,t}(I-\delta) - \delta \qquad \text{(C.8)}$$

which contradicts (C.6) under the assumption $q^*_{f,t}(I) > q^*_{f,t}(I-\delta)$. Hence, $q^*_{f,t}(I) \leq q^*_{f,t}(I-\delta)$. We can show that $q^*_{s,t}(I) \leq q^*_{s,t}(I-\delta)$ in a similar fashion.

Next, we show that $q^*_{j,t}(I-\delta) - q^*_{j,t}(I) \leq \delta$, $j = f, s$. First, assume that $q^*_{f,t}(I-\delta) - \delta > q^*_{f,t}(I)$. Note that under this assumption $\phi_f(I; q^*_{f,t}(I), q^*_{s,t}(I)) - \phi_f(I-\delta; q^*_{f,t}(I-\delta) \geq 0$. Further, by the convexity of $H(\cdot)$ and $V_{t+1}(\cdot)$, the assumption implies:

$$\begin{aligned}
&\mathbb{E}H'(I + q^*_{f,t}(I) - D) \leq \mathbb{E}H'(I - \delta + q^*_{f,t}(I-\delta) - D)\\
&\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) - D)\} \leq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) - D)\}\\
&\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + k_s - D)\} \leq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) + k_s - D)\}
\end{aligned}$$

Then, for $\phi_f(I; q^*_{f,t}(I), q^*_{s,t}(I)) - \phi_f(I-\delta; q^*_{f,t}(I-\delta) \geq 0$ to hold, the following condition must be satisfied:

$$\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + q^*_{s,t}(I) - D) \geq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - D)$$

which implies

$$q^*_{f,t}(I) + q^*_{s,t}(I) \geq q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - \delta \tag{C.9}$$

Since we assumed that $q^*_{f,t}(I-\delta)-\delta > q^*_{f,t}(I)$, (C.9) requires $q^*_{s,t}(I) > q^*_{s,t}(I-\delta)$, a contradiction because $q^*_{s,t}(I) \leq q^*_{s,t}(I-\delta)$. Hence, $q^*_{f,t}(I-\delta) \leq q^*_{f,t}(I) + \delta$. We can show that $q^*_{s,t}(I-\delta) \leq q^*_{s,t}(I) + \delta$ in a similar fashion.

Finally, we show that $q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) \geq q^*_{f,t}(I) + q^*_{s,t}(I) + \delta$. First consider $q^*_{f,t}(I) > 0$. Note that since $q^*_{f,t}(I-\delta) - \delta \leq q^*_{f,t}(I)$, the following relations hold:

$$\mathbb{E}H'(I + q^*_{f,t}(I) - D) - \mathbb{E}H'(I - \delta + q^*_{f,t}(I-\delta) - D) \geq 0$$
$$\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) - D) \geq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) - D)$$
$$\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + k_s - D) \geq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) + k_s - D)$$

Hence, in order to satisfy $\phi_f(I; q^*_{f,t}(I), q^*_{s,t}(I)) = \phi_f(I-\delta; q^*_{f,t}(I-\delta), q^*_{s,t}(I-\delta))$, the following must hold:

$$\mathbb{E}V'_{t+1}(I + q^*_{f,t}(I) + q^*_{s,t}(I) - D) \leq \mathbb{E}V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - D)$$

which implies

$$q^*_{f,t}(I) + q^*_{s,t}(I) \leq q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - \delta$$

Now consider $q^*_{f,t}(I) = 0$. Since $I \leq max\{I_{f,t}, I_{s,t}\}$, we have $q^*_{s,t}(I) > 0$. By the first order condition for $q_{s,t}(I)$

$$\begin{aligned}\alpha \mathbb{E} V'_{t+1}(I + q^*_{s,t}(I) - D) &= \mathbb{E} V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) \wedge k_f + q^*_{s,t}(I-\delta) - D) \\ &\leq \mathbb{E} V'_{t+1}(I - \delta + q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - D)\end{aligned}$$

which implies

$$q^*_{s,t}(I) = q^*_{s,t}(I) + q^*_{f,t}(I) \leq q^*_{f,t}(I-\delta) + q^*_{s,t}(I-\delta) - \delta$$

concluding the proof. □

Proof of Proposition 4.4.3.

We first prove part i). Suppose the result is not true and thus $I_{f,t} < I_{s,t}$. We have $q^*_{f,t}(I_{f,t}) = 0$ and $q^*_{s,t}(I_{f,t}) > 0$. Since $I_{f,t}$ is the re-order point, we have from the first-order condition of q_f that

$$\begin{aligned}&\phi_{f,t}(I_{f,t}, 0, q^*_{s,t}(I_{f,t})) = \\ &c_f - p + (h+p)F(I_{f,t}) + \alpha \mathbb{E} V'_{t+1}(I_{f,t} + q^*_{s,t}(I_{f,t}) \wedge k_s - D) = 0,\end{aligned} \tag{C.10}$$

where $F(\cdot)$ is the distribution function for the demand D. By the first-order condition of $q_{s,t}(I)$, we have

$$\phi_{s,t}(I_{f,t}, 0, q^*_{s,t}(I_{f,t})) = c_s + \alpha \mathbb{E} V'_{t+1}(I_{f,t} + q^*_{s,t}(I_{f,t}) - D) = 0.$$

Furthermore, we note that

$$\begin{aligned}
& \alpha\mathbb{E}V'_{t+1}(I_{f,t}+q^*_{s,t}(I_{f,t})\wedge k_s - D) \\
= \; & \alpha\mathbb{E}_{k_s}\min\{\mathbb{E}_D V'_{t+1}(I_{f,t}+q^*_{s,t}(I_{f,t})-D), \mathbb{E}_D V'_{t+1}(I_{f,t}+k_s-D)\} \\
= \; & \alpha\mathbb{E}_{k_s}\min\{-c_s/\alpha, \mathbb{E}_D V'_{t+1}(I_{f,t}+k_s-D)\} \\
\leq \; & -c_s,
\end{aligned}$$

where $\mathbb{E}_X$ represents the expectation with respect to the random variable X. Substituting the above relation in (C.10),

$$c_f - p + (h+p)F(I_{f,t}) - c_s \geq 0$$

Hence, we deduce that

$$F(I_{f,t}) \geq \frac{p - c_f + c_s}{h+p}.$$

Since $I_{f,t} < I_{s,t}$, $I_{f,t}$ must be bounded. We must have $F(I_{f,t}) \leq 1$. This, in turn, implies that $c_s \leq c_f + h$, which contradicts the condition in part i). Hence, we must have $I_{f,t} \geq I_{s,t}$.

To see part ii) suppose the result is not true and thus $I_{f,t} > I_{s,t}$. We have $q^*_{s,t}(I_{s,t}) = 0$ and $q^*_{f,t}(I_{s,t}) > 0$. Since $I_{s,t}$ is the re-order point, we have from the first-order condition of $q_{s,t}(I)$ that

$$\phi_{s,t}(I_{s,t}, q^*_{f,t}(I_{s,t}), 0) = c_s + \alpha\mathbb{E}V'_{t+1}(I_{s,t}+q^*_{f,t}(I_{s,t})\wedge k_f - D) = 0. \quad \text{(C.11)}$$

By the first-order condition of $q_{f,t}(I)$, we have

$$\begin{aligned}
& \phi_{f,t}(I_{s,t}, q^*_{f,t}(I_{s,t}), 0) = \\
& c_f - p + (h+p)F(I_{s,t}+q^*_{f,t}(I_{s,t})) + \alpha\mathbb{E}V'_{t+1}(I_{s,t}+q^*_{f,t}(I_{s,t}) - D) = 0,
\end{aligned}$$

Furthermore, we note that

$$\begin{aligned}
&\alpha \mathbb{E} V'_{t+1}(I_{s,t} + q^*_{f,t}(I_{s,t}) \wedge k_f - D) \\
= \;& \mathbb{E}_{k_f} \min\{\alpha \mathbb{E}_D V'_{t+1}(I_{s,t} + q^*_{f,t}(I_{s,t}) - D), \alpha \mathbb{E}_D V'_{t+1}(I_{s,t} + k_f - D)\} \\
= \;& \mathbb{E}_{k_f} \min\{-c_f + p - (h+p)F(I_{s,t} + q^*_{f,t}(I_{s,t})), \alpha \mathbb{E}_D V'_{t+1}(I_{s,t} + k_f - D)\} \\
\leq \;& -c_f + p - (h+p)F(I_{s,t} + q^*_{f,t}(I_{s,t}))
\end{aligned}$$

Substituting the above relation in (C.11),

$$c_s - c_f + p - (h+p)F(I_{s,t} + q^*_{f,t}(I_{s,t})) \geq 0$$

Hence, we deduce that

$$F(I_{s,t} + q^*_{f,t}(I_{s,t})) \leq \frac{c_s - c_f + p}{h+p}.$$

By definition, we must have $F(I_{s,t} + q^*_{f,t}(I_{s,t})) \geq 0$. This, in turn, implies that $c_f \leq p + c_s$, which contradicts the condition in part ii). Hence, we must have $I_{f,t} \leq I_{s,t}$. □

Remark C.0.1. We consider the case when the fast supplier has an L-period leadtime and the slow one has an $L+1$-period leadtime.To ensure tractability, we make a simplifying assumption that the supplier capacities are observed one period after the corresponding orders are placed. In other words, $k_{f,t}$ and $k_{s,t}$ are realized at the end of period t when the demand D_t materializes. In this case, the total available inventory at the beginning of period t before orders $q_{f,t}$ and $q_{s,t}$ are issued, i.e., the pre-order inventory position, is

$$y_t = I_t + \sum_{n=t-L+1}^{t-1} q_{f,n} \wedge k_{f,n} + \sum_{n=t-L}^{t-1} q_{s,n} \wedge k_{s,n}.$$

Note that $k_{f,n}$ and $k_{s,n}$, $n < t$, are constants in period t on account of our assumption. It is easily seen that y_t follows the following dynamics:

$$y_{t+1} = y_t + q_{f,t} \wedge k_{f,t} + q_{s,t} \wedge k_{s,t} - D_t.$$

The amount of net inventory at the beginning of period $t+L$ before the order $q_{f,t}$ arrives is

$$I_{t+L} = y_t - D^{(L)},$$

where $D^{(L)}$ is the Lth-fold convolution of D. Thus, the objective of the problem becomes

$$\begin{aligned} J_t(y; q_{f,t}, q_{s,t}) &= c_f \mathbb{E}(q_{f,t} \wedge k_f) + c_s \mathbb{E}(q_{s,t} \wedge k_s) + \alpha^{L-1} \mathbb{E} H(y + q_{f,t} \wedge k_f - D^{(L)}) \\ &\quad + \alpha^L \mathbb{E} V_{t+L}(y + q_{f,t} \wedge k_f + q_{s,t} \wedge k_s - D^{(L)}). \end{aligned}$$

It is easily shown that all the results derived earlier for the case $L = 0$ can be generalized to any positive L.

Bibliography

Abdel-Hamid, T. (1984), The Dynamics of Software Development Project Management: An Integrative System Dynamics Perspective, PhD thesis, MIT, Cambridge, MA.

Abdel-Hamid, T. and Madnick, S. E. (1991), *Software Project Dynamics: An Integrated Approach*, Prentice-Hall, Upper Saddle River, NJ.

Allen, D. (2006), 'First look: Sony PS3 versus Nintendo Wii'. Retrieved November 14, 2006, from PC World. `http://www.pcworld.com/`.

Allon, G. and Van Mieghem, J. A. (2010), 'Global dual sourcing: Tailored base-surge allocation to near- and offshore production.', *Management Science* **56**(1), 110–124.

American Petroleum Institute (2005), 'Basic petroleum data book: Petroleum industry statistics, volume xxv, number 2'.

American Wind Energy Association (2006), 'Economies of wind energy, wind energy factsheets'. `www.awea.org`.

Anderson, E. G. and Parker, G. G. (2002), 'The effect of learning on the make/buy decision', *Production and Operations Management* **11**(3), 313–339.

Anderson, S. and Coate, S. (2005), 'Market provision of broadcasting: A welfare analysis', *Review of Economic Studies* **72**, 947–972.

Anupindi, R. and Akella, R. (1993), 'Diversification under supply uncertainty', *Management Science* **39**(8), 944–963.

Argote, L. (1999), *Organizational Learning: Creating, Retaining, and Transferring Knowledge*, Kluwer Academic, Boston, MA.

Armstrong, M. (2002), The theory of access pricing and interconnection, *in* M. Cave, S. Majumdar and I. Vogelsang, eds, 'Handbook of Telecommunications Economics: Vol.I.', North Holland, Amsterdam.

Armstrong, M. (2006), 'Competition in two-sided markets', *RAND Journal of Economics* **37**(3), 668–691.

Armstrong, M. and Wright, J. (2007), 'Two-sided markets, competitive bottlenecks and exclusive contracts', *Economic Theory* **32**, 353–380.

Arthur, B. (1989), 'Competing technologies, increasing returns, and lock-in by historical events', *Economic Journal* **99**(March), 116–131.

Barankin, E. (1961), 'A delivery-lag inventory model with an emergency provision (the single-period case)', *Naval Research Logistics Quarterly* **8**(3), 285–311.

Bertsekas, D. (2005), *Dynamic Programming and Optimal Control, vol. I.*, Athena Scientific, Nashua, NH.

Bhargava, H. K. and Choudhary, V. (2004), 'Economics of an information intermediary with aggregation benefits', *Information Systems Research* **15**(1), 22–36.

Black, F. and Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**(3), 637–654.

Borison, A. (2005), 'Real options analysis: Where are the emperors clothes?', *Journal of Applied Corporate Finance* **17**(2), 17–32.

Boudreau, K. and Hagiu, A. (2008), Platform rules: Multi-sided platforms as regulators. SSRN: `http://papers.ssrn.com/abstract=1269966`.

BP Global (2006), 'Statistical review of world energy 2006'.

Brandao, L. E. and Dyer, J. S. (2005), 'Decision analysis and real options: A discrete time approach to real option valuation', *Annals of Operations Research* **135**(1), 21–39.

Brandao, L. E., Dyer, J. S. and Hahn, W. (2005), 'Using binomial decision trees to solve real-option valuation problems', *Decision Analysis* **2**(2), 69–88.

Brandenburger, A. M. and Nalebuff, B. J. (1996), *Co-opetition*, Doubleday, New York.

Burke, G. J., Carrillo, J. E. and Vakharia, A. J. (2007), 'Single versus multiple supplier sourcing strategies', *European Journal of Operational Research* **182**, 95–112.

Caillaud, B. and Jullien, B. (2003), 'Chicken and egg: Competition among intermediation service providers', *RAND Journal of Economics* **34**(2), 309–328.

Capozza, D. and Li, Y. (1994), 'The intensity and timing of investment: The case of land', *American Economic Review* **84**(4), 889–904.

Casual Games Association (2007), 'Casual games market report'.

Chao, X., Chen, H. and Zheng, S. (2008), 'Joint replenishment and pricing decisions in inventory systems with stochastically dependent supply capacity', *European Journal of Operational Research* **191**, 142–155.

Cheng, J. (2007), 'Report: Video game spending to surpass music spending this year'. Retrieved June 23, 2007, from Ars Technica- the art of technology. `http://arstechnica.com/index.ars`.

Chiang, C. and Benton, W. C. (1994), 'Sole sourcing versus dual sourcing under stochastic demands and lead times.', *Naval Research Logistics* **41**(5), 609–624.

Ciarallo, F. W., Akella, R. and Morton, T. E. (1994), 'A periodic review production planning model with uncertain capacity and uncertain demand-optimality of extended myopic policies', *Management Science* **40**(3), 320–332.

Clemen, R. (1997), *Making Hard Decisions: An Introduction to Decision Analysis*, Duxbury Press, Belmont, CA.

Copeland, T. and Antikarov, V. (2001), *Real Options: A Practiotioner's Guide*, Texere, New York.

Dada, M., Petruzzi, N. C. and Schwarz, L. B. (2007), 'A newsvendor's procuremnt problem when suppliers are unreliable.', *Manufacturing and Service Operations Management* **9**(1), 9–32.

Daniel, K. H. (1963), A delivery-lag inventory model with emergency, *in* H. E. Scarf, D. M. Gilford and M. W. Shelly, eds, 'Multistage Inventory Models and Techniques', Stanford University Press, California.

Davidsen, P. I., Sterman, J. D. and Richardson, G. P. (1990), 'A petroleum life cycle model for the united states with endogenous technology, exploration, recovery, and demand', *System Dynamics Review* **6**(1), 66–93.

Dixit, A. and Pindyck, R. (2001), *Investment under Uncertainty*, Princeton University Press, Princeton, NJ.

Dutka, B. (2009), 'Do multiplatform games affect console sales?'. Retrieved June 4, 2009, from PSXExtreme. `http://www.psxextreme.com/feature/410.html`.

Economides, N. and Katsamakas, E. (2006), 'Two-sided competition of proprietary vs. open source technology platforms and the implications for the software industry', *Management Science* **52**(7), 1057–1071.

EIA (2007), Annual Energy Outlook 2007, Technical Report No: DOE/EIA-0383(2007).

Eisenmann, T. R., Parker, G. and Alstyne, M. W. V. (2006), 'Strategies for two-sided markets', *Harvard Business Review* **October**.

Eisenmann, T. R., Parker, G. and Alstyne, M. W. V. (2007), 'Platform envelopment'. HBS Technology and Operations Mgt. Unit Research Paper No. 07-104; MIT Sloan Research Paper No. 4681-08. Available at SSRN: `http://ssrn.com/abstract=996852`.

EMA (2008), 'The Entertainment Merchant Association Annual Report'.

Energy Information Administration (EIA) (2006), 'Department of energy, us data'.

Federgruen, A. and Yang, N. (2008), 'Selecting a portfolio of suppliers under demand and supply risks', *Operations Research* **56**(4), 916–936.

Federgruen, A. and Yang, N. (2009), 'Optimal supply diversification under general supply risks', *Operations Research* **57**(6), 1451–1468.

Federgruen, A. and Zipkin, P. (1986), 'An inventory model with limited production capacity and uncertain demands II. the discounted-cost criterion', *Mathematics of Operations Research* **11**(2), 208–215.

Feng, Q., Sethi, S., Yan, H. and Zhang, H. (2006), 'Are base-stock policies optimal in inventory problems with multiple delivery modes?', *Operations Research* **54**(4), 801–807.

Ford, A. (1997), 'System dynamics and the electric power industry', *System Dynamics Review* **13**(1), 57–85.

Ford, A. (2006), Simulating the impact of carbon market on the electricity system in the western u.s.a., *in* 'Proceedings of the 24th International Conference of the System Dynamics Society', Nijmegen, Netherlands.

Ford, D. and Bhargav, S. (2006), 'Project management quality and the value of flexible strategies', *Engineering, Construction and Architectural Management* **13**(3), 275–289.

Ford, D. and Sobek, S. (2005), 'Adapting real options to new product development by modeling the Second Toyota Paradox', *IEEE Transactions on Engineering Management* **52**(2), 175–185.

Forrester, J. (1961), *Industrial Dynamics*, MIT Press, Cambridge, MA.

Forrester, J. (1975), The impact of feedback control systems on the management sciences, *in* 'Collected Papers of Jay W. Forrester', MIT Press, Cambridge, MA.

Foster, P. and Ahmad, M. (2008), 'Computer games to out-sell music and video'. Retrieved November 5, 2008, from Times Online. `http://business.timesonline.co.uk/tol/business/`.

Fukuda, Y. (1964), 'Optimal policies for the inventory problem with negotiable lead time', *Management Science* **10**(4), 690–708.

Gawer, A. and Cusumano, M. A. (2002), *Platform leadership: How Intel, Microsoft, and Cisco drive industry innovation*, Harvard Business School Publishing, Boston.

Gerchak, Y. and Parlar, M. (1990), 'Yield randomness, cost tradeoffs, and diversification in the eoq model', *Naval Research Logistics* **37**(3), 341–354.

Green, P. E. and Srinivasan, V. (1990), 'Conjoint analysis in marketing: New developments with implications for research and practice', *Journal of Marketing* **54**(4), 3–19.

Hagiu, A. (2006), 'Pricing and commitment by two-sided platforms', *RAND Journal of Economics* **37**(3), 720–737.

Hagiu, A. and Lee., R. (2007), Exclusivity and control. Strategy Unit Working Paper No. 08-009. Available at SSRN: `http://ssrn.com/abstract=1010669`.

Hahn, W. J. and Dyer, J. S. (2008), 'Discrete time modeling of mean-reverting stochastic processes for real option valuation', *European Journal of Operational Research* **184**, 534–548.

Henig, M. and Gerchak, Y. (1990), 'The structure of periodic review policies in the presence of random yield.', *Operations Research* **38**, 634–643.

Heyman, D. P. and Sobel, M. J. (1984), *Stochastic Models in Operations Research, Vol. II: Stochastic Optimization*, McGraw-Hill, New York.

Hull, J. (2006), *Options, Futures and Other Derivatives*, Prentice Hall, New Jersey.

International Energy Agency (IEA) (2006), 'World energy outlook 2006'.

Johnson, S. T., Taylor, T. and Ford, D. (2006), Using system dynamics to extend real options use: Insights from the oil and gas industry, *in* 'Proceedings of the System Dynamics Conference', Nijmegen, Netherlands.

Junginger, M., Faaij, A. and Turkenberg, W. C. (2005), 'Global experience curves for wind farms', *Energy Policy* **33**, 133–150.

Kaiser, U. and Wright, J. (2006), 'Price structure in two-sided markets: Evidence from the magazine industry', *International Journal of Industrial Organization* **24**, 1–28.

Kane, Y. I. and Wakabayashi, D. (2009), 'Nintendo looks outside the box: Videogame maker seeks third-party publishers to expand wii library'. Retrieved June 4, 2009, from Wall Street Journal. `http://online.wsj.com/article/SB124338239503456555.html`.

Kane, Y. I. and Wingfield, N. (2007), 'Nintendo plays it a wii bit cautious', *Wall Street Journal* . December 7, 2007.

Katz, M. L. and Shapiro, C. (1985), 'Network externalities, competition, and compatibility', *American Economic Review* **75**(3), 424–440.

Katz, M. L. and Shapiro, C. (1994), 'Systems competition and network effects', *Journal of Economic Perspectives* **8**(2), 93–115.

Kogut, B. and Kulatilaka, N. (1994), 'Operating flexibility, global manufacturing, and the option value of a multinational network.', *Management Science* **40**(1), 123–139.

Kouvelis, P. and Li, J. (2010), 'Offshore outsourcing, yield uncertainty and contingency responses.', *Working Paper* .

Krishnan, V. and Ulrich, K. T. (2001), 'Product development decisions: A review of the literature', *Management Science* **47**(1), 1–21.

Lawson, D. G. and Porteus, E. L. (2000), 'Multistage inventory management with expediting', *Operations Research* **48**(6), 878–893.

Leheng, Q. (2006), 'Wii development costs a quarter to half compared to PS3/360'. Retrieved August 11, 2006, from TechGate. `http://tech.commongate.com/`.

Liebowitz, S. and Margolis, S. E. (1994), 'Network externality: An uncommon tragedy', *Journal of Economic Perspectives* **8**(2), 133–150.

Lyneis, J. M. and Ford, D. N. (2007), 'System dynamics applied to project management: A survey, assessment, and directions for future research', *System Dynamics Review* **23**(2), 157–189.

McDonald, R. and Siegel, D. (1986), 'The value of waiting to invest', *Quarterly Journal of Economics* **101**, 701–727.

McIntyre, T. (2008), 'Nintendo is staging wii shortage to feed popularity hype and inflate stock'. Retrieved October 7, 2010 `http://tech.blorge.com/Structure:-shortage-to-feed-popularity-hype-and-inflate-stock/`.

Money, M. (2007), 'Video-game sales overtaking music'. Retrieved June 26, 2007, from MSN Money. `http://moneycentral.msn.com/home.asp`.

Morecroft, J. D. W. and van der Heijden, K. A. J. M. (1992), 'Modeling the oil producers: Capturing oil industry knowledge in a behavioral simulation model', *European Journal of Operational Research* **59**(1), 102–122.

Musial, W. and Butterfield, S. (2004), Future of offshore wind energy in the united states, nrel/cp-500-36313., *in* 'Energy Ocean Proceedings', Palm Beach, FL.

Mussa, M. and Rosen, S. (1978), 'Monopoly and product quality', *Journal of Economic Theory* **18**, 301–317.

Myers, S. (1987), 'Finance theory and financial strategy', *Midland Corporate Finance Journal* **5**, 6–13.

Naill, R. (1992), 'A system dynamics model for national energy policy planning', *System Dynamics Review* **8**(1), 1–19.

National Renewable Energy Laboratory (2006), 'Concentrating solar power research, 2006'. http://www.nrel.gov/csp/publications.html.

Nau, R. and McCardle, K. (1991), 'Arbitrage, rationality and equilibrium', *Theory and Decision* **31**(2), 199–240.

Northwest Power and Conservation Council (2007), 'Biennial monitoring report on the fifth power plan'. Retrieved on June 10, 2007. http://www.nwcouncil.org/library/2007/2007-4.htm.

Nutt, C. (2008), 'Sony's jack tretton: The full E3 roundtable report.'. Retrieved June 8, 2009 from Gamasutra.com. http://www.gamasutra.com/php-bin/news_index.php?story=19447.

Osgood, N. (2005), Combining system dynamics and decision analysis for rapid strategy selection, *in* 'Proceedings of the 23rd International Conference of System Dynamics Society', Boston, US.

Osgood, N. and Kaufman, G. (2003), A hybrid model architecture for strategic renewable resource planning, *in* 'Proceedings of the 21st International Conference of System Dynamics Society', New York City, US.

Paddock, J., Siegel, D. R. and Smith, J. L. (1988), 'Option valuation of claims on real assets: The case of offshore petroleum leases', *Quarterly Journal of Economics* **103**, 479–508.

Papalambros, P. Y. (1995), 'Optimal-design of mechanical-engineering systems', *J. Mechanical Design* **117**(June), 48–54.

Parker, G. and Alstyne, M. V. (2000), Information complements, substitutes, and strategic product design. mimeo, Tulane University and University of Michigan.

Parker, G. and Alstyne, M. V. (2005), 'Two-sided network effects: A theory of information product design', *Management Science* **51**(10), 1494–1504.

Parker, G. and Alstyne, M. V. (2008), Innovation, openness and platform control. mimeo, Tulane University and MIT.

Parlar, M. and Wang, D. (1993), 'Diversification under yield randomness in inventory models', *European Journal of Operational Research* **66**(1), 52–64.

Pindyck, R. (1991), 'Irreversibility, uncertainty, and investment', *American Economic Association* **29**(3), 1110–1148.

Pugh, E. W. (1984), *Memories that Shaped an Industry, Decisions Leading to IBM Systems/360*, MIT Press, Cambridge, MA.

Reisinger, D. (2009), 'Why first-party video games matter.'. Retrieved June 8, 2009. `http://news.cnet.com/8301-13506_3-10236513-17.html`.

Roberts, E. B. (1974), A simple model of R & D project dynamics, *in* E. B. Roberts, ed., 'Managerial Applications of System Dynamics', Productivity Press, Cambridge, MA.

Rochet, J. C. and Tirole, J. (2003), 'Platform competition in two-sided markets', *J. Eur. Econom. Assoc.* **1**(4), 990–1029.

Rochet, J. C. and Tirole, J. (2006), 'Two-sided markets: A progress report', *RAND Journal of Economics* **37**(3), 645–667.

Scheller-Wolf, A., Veeraraghavan, S. and Houtum, G. V. (2005), Inventory models with expedited ordering: Single index policies. Working Paper, Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA.

Sethi, S. P., Yan, H. and Zhang, H. (2003), 'Inventory models with fixed costs, multiple delivery modes and forecast updates.', *Operations Research* **51**(2), 321–328.

Shane, S. A. and Ulrich, K. T. (2004), 'Technological innovation, product development, and entrepreneurship in management science', *Management Science* **50**(2), 133–144.

Sheffield, B. (2008), 'The graphics plateau', *Game Developer* **15**(7).

Sheopuri, A., Janakiraman, G. and Seshadri, S. (2010), 'New policies for the stochastic inventory control problem with two supply modes.', *Operations Research* **58**(3), 734–745.

Sinclair, B. (2006), 'Wii dev costs fraction of PS3's, 360's'. Retrieved May 5, 2006, from Gamespot. `http://www.gamespot.com/`.

Smith, J. (1999), ‘Much ado about options?’, *Decision Analysis Newsletter* **18**(2), 4–8.

Smith, J. (2005), ‘Alternative approaches for solving real-options problems’, *Decision Analysis* **2**(2), 89–102.

Smith, J. and McCardle, K. (1998), ‘Valuing oil properties: Integrating option pricing and decision analysis approaches’, *Operations Research* **46**(2), 198–217.

Smith, J. and Nau, R. (1995), ‘Valuing risky projects: Option pricing theory and decision analysis’, *Management Science* **14**(5), 795–816.

Sterman, J. (1981), The Energy Transition and the Economy: A System Dynamics Approach, PhD thesis, MIT, Cambridge, MA.

Sterman, J. (2000), *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw-Hill/Irwin, New York.

Sterman, J. D. and Richardson, G. P. (1985), ‘An experiment to evaluate methods for estimating fossil fuel resources’, *Journal of Forecasting* **4**(2), 197–226.

Sterman, J. D., Richardson, G. P. and Davidsen, P. I. (1988), ‘Modeling the estimation of petroleum resources in the united states’, *Technological Forecasting and Social Change* **33**(3), 219–249.

Swaminathan, J. M. and Shanthikumar, J. G. (1999), 'Supplier diversification: effect of discrete demand.', *Operations Research Letters* **24**, 213–221.

Teisberg, E. O. (1995), Methods for evaluating capital investment decisions under uncertainty. real options, *in* L. Trigeorgis, ed., 'Capital Investment: New Contributions', Praeger Publishing, Westport, CT.

Thomke, S. (1999), 'Project dreamcast: Serious play at Sega Enterproses Ltd. (A)'. Harvard Business School, case study N9-600-028.

Tomlin, B. (2006), 'On the value of mitigation and contingency strategies for managing supply chain disruption risks', *Management Science* **52**(5), 639–657.

Trigeorgis, L. (1991), 'A log-transformed binomial numerical analysis method for valuing complex multi-option investments', *Journal of Financial and Quantitative Analysis* **26**(3), 309–326.

Trigeorgis, L. and Mason, S. P. (1987), 'Valuing managerial flexibility', *Midland Corporate Finance Journal* **5**, 14–21.

van der Rhee B., Schmidt, G. and Tsai., W. (2009), Steepen, maintain, or flatten the performance treadmill? Working Paper, Nyenrode Business University, Breukelen, The Netherlands., David Eccles School of Business, University of Utah, Salt Lake City, UT.

Veeraraghavan, S. and Scheller-Wolf, A. (2008), 'Now or later: A simple policy for effective dual sourcing in capacitated systems.', *Operations Research* **56**(4), 850–864.

VG Chartz Game Database (2009). Retrieved 24 January 2009. `http://vgchartz.com/games/`.

VG Chartz Home Page (2009). Retrieved 21 January 2009. `http://vgchartz.com/`.

Vogstad, K. (2004), A System Dynamics Analysis of the Nordic Electricity Market: The Transition from Fossil Fuelled Toward a Renewable Supply within a Liberalized Electricity Market, PhD thesis, Norwegian University of Science and Technology.

Wang, T. and Dyer, J. (2009), 'Valuing multifactor real options using a binomial tree.', *Decision Analysis* **7**(2), 185–195.

Wang, Y. and Gerchak, M. (1996), 'Periodic review production models with variable capacity, random yield, and uncertain demand', *Management Science* **42**(1), 130–137.

Wang, Y., Gilland, W. and Tomlin, B. (2010), 'Mitigating supply risk: Dual sourcing or process improvement?', *Manufacturing and Service Operations Management* **12**(3), 489–510.

Whittmore, A. and Saunders, S. C. (1977), 'Optimal inventory under stochastic demand with two supply options.', *SIAM Journal on Applied Mathematics* **12**(2), 293–305.

Wingfield, N. and Guth, R. A. (2005), 'Why shortages of hot gifts endure as a christmas ritual', *The Wall Street Journal* **December 2,2005**.

Yang, J., Qi, X. and Xia, Y. (2005), 'A production-inventory system with markovian capacity and outsourcing option', *Operations Research* **53**(2), 328–349.

Yano, C. and Lee, H. (1995), 'Lot sizing with random yields: A review', *Operations Research* **43**(2), 311–334.

Yazlali, O. and Erhun, F. (2009), 'Dual-supply inventory problem with capacity limits on order sizes and unrestricted ordering costs', *IIE Transactions* **41**(8), 716–729.

Zhu, F. and Iansiti, M. (2009), Dynamics of platform competition: Exploring the role of installed base, platform quality and consumer expectations. HBS Working Paper 08-031, Harvard Business School, Boston, MA.

Vita

Burcu Tan was born in Izmir, Turkey. She received her Bachelor of Science and Master of Science degrees in Industrial Engineering from Bogazici University in Turkey. She joined the PhD program at McCombs School of Business in 2005 to study Operations Management. Her research and teaching interests include operations and risk management interface, two-sided markets, supply chain management, and system dynamics.

Permanent address: 1906 Pearl St. Apt 304
Austin, Texas 78705

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.